

Universitetet i Tartu
Institutt for germansk, romansk og slavisk filologi
Avdeling for skandinavistikk
Norsk språk og litteratur

**„*Frosken, hvor er du?*“ – syntaktisk kompleksitet
i skriftlige fortellinger av estiske innlærere
av norsk som fremmedspråk.**

Masteroppgave

Karoliine Piksarv

Veileder: Anu Laanemets, PhD

Innholdsfortegnelse

1 INNLEDNING	3
2 TEORI OG METODE	5
2.1 Forskningsfeltet CAF (<i>Complexity, Accuracy, Fluency</i>).....	5
2.2 Syntaktisk kompleksitet.....	6
2.3 Forklaringer for noe sentrale begrep	9
2.4 Målestokk for syntaktisk kompleksitet.....	15
3 MATERIALET OG INFORMANTER.....	22
4 ANALYSE OG RESULTATER.....	25
4.1 Nybegynnergruppen (Gruppe A).....	25
4.1.1 Generelt om tekstene	25
4.1.2 Hvilke strukturer brukes?	27
4.1.3 Syntaktisk korrekthet.....	30
4.2 Viderekommende gruppen (Gruppe B)	36
4.2.1 Generelt om tekstene	36
4.2.2 Hvilke strukturer brukes?	36
4.2.3 Syntaktisk korrekthet.....	40
4.3 Sammenlikning av alle grupper (inkludert gruppe C: morsmålstalere)	45
4.3.1 Generelt om tekstene	45
4.3.2 Hvilke strukturer brukes?	48
4.3.3 Analysen av de lengste syntaktiske enheter	50
4.3.4 Syntaktisk korrekthet.....	55
5 DISKUSJON OG KONKLUSJON	58
6 EESTIKEELNE RESÜMEE	63
LITTERATURLISTE	69
Tillegg 1: Bildene fra Frog, where are you? (Mayer 1969)	71
Tillegg 2: Lihtlitsents	75

1 Innledning

Foreliggende oppgave har til hensikt å kartlegge norsk innlærerspråk til studenter med estisk som morsmål ved å analysere studentenes skriftlige produksjon. Fokuset ligger ved tilegnelsen av de forskjellige syntaktiske strukturer og den syntaktiske kompleksiteten. Formålet er å sammenlikne den syntaktiske kompleksiteten i studentenes tekster på to forskjellige utviklingsnivåer – nybegynnere og mellomnivå, samt å se på hvordan deres mellomspråk forholder seg til målspråket, det vil si tekstene til morsmålstalende informanter.

Gjennom analysen vil jeg finne svar på følgende spørsmål:

- 1) Hva som kjennetegner syntaksen på de respektive utviklingsnivåene og i hvilke trekk er den forskjellig fra språkbruket til morsmålstalere?
- 2) Hvorvidt brukes komplekse syntaktiske strukturer på de forskjellige utviklingsnivåene? Gjennom hvilke trekk kommer den syntaktiske kompleksiteten til syne? Hvilke typer av strukturer studentene bruker?
- 3) Hvilke av strukturene har studentene tilegnet seg framgangsrikt og hvilke ikke? Hvilke tendenser finnes det med hensyn til strukturer som studentene har problemer med?
- 4) Hvordan kan likhetene og forskjellene mellom gruppene forklares?

Målet med undersøkelsen er å kartlegge hvilke syntaktiske strukturer studentene bruker, hvilke de allerede mestrer og hvilke de har fremdeles problemer med. Analysen dreier seg hovedsakelig om den syntaktiske kompleksiteten på ytringsnivå og utvides ikke over ytringsgrenser. Samt har jeg utelatt mindre analyseenheter på frasenivå og setningsdeler når det gjelder den kvantitative analysen på grunn av masteroppgavens begrensede omfang. Likevel vil disse nivåer bli kommentert i forbindelse med den kvalitative analysen.

Undersøkelsen skal også gi innblikk i hvordan de voksne estiske innlærere (studenter ved universitetet) tilegner seg norsk syntaks. Samtidig som resultatene skal komplettere vår forståelse av syntaktisk kompleksitet innenfor andrespråkstilegnelsen og forskningsfeltet CAF (*Complexity, Accuracy, Fluency*), håper jeg også å gi mitt bidrag til den pedagogiske diskusjonen og skape et utgangspunkt for overveielser over innholdet og rekkefølgen av

undervisningen av norsk som fremmedspråk til estiske studenter. Resultatene kan utvides til videre pedagogiske diskusjoner om hvordan man kan støtte studentenes utvikling på best mulig måte og hvilke områder i språktilegnelsen man skal legge mest vekt på.

Oppgaven består av 5 kapitler. I kapittel 1 har jeg presentert oppgavens mål og forskningsspørsmål. I kapittel 2 gir jeg en oversikt over teoriene som er relevante for analysen, med særlig vekt på forskningsfeltet CAF og forskning av kompleksitet i andrespråkstilegnelsen. Kapitlet inneholder også definisjonene for de mest sentrale begrep slik som jeg har brukt dem i min analyse, samt forklaringen og begrunnelser av måleenhetene som analysen bygges på. I kapittel 3 legger jeg fram materialet som analysene baseres på og kommenterer på hvordan dette ble innsamlet. Kapittel 4 presenterer analyseresultatene og diskuterer funnene nærmere. Resultatene blir framlagt separat for de to studentgruppene, deretter foretar jeg en sammenlikning av gruppene seg imellom, samt sammenlikning med morsmålsgruppen. Analysen vil bli illustrert med tabeller over resultatene og eksempler fra tekstene. Kapittel 6 skal bestå av diskusjonen av de mest sentrale funn og konklusjoner. Til slutt kommer det et sammendrag på estisk.

2 Teori og metode

2.1 Forskningsfeltet CAF (*Complexity, Accuracy, Fluency*)

Et av de sentrale spørsmålene i andrespråksforskningen (og spesielt innenfor anvendt lingvistikk) har vært: Hva som kjennetegner en mer eller mindre «dyktig» språkbruker (enten det er en andrespråks- eller en innfødt språkbruker)? Hvilke trekk må vi se på for å beskrive språkutviklingen? Innenfor anvendt lingvistikk snakker man om multidimensjonal språkutvikling og har delt disse kjennetegn inn i 3 større kategorier: kompleksitet (*complexity*), nøyaktighet (*accuracy*) og flyt (*fluency*), dermed kalles forskningsfeltet for kort CAF ut fra de engelske termene. CAF brukes både for å formulere indikatorer for vurdering av studentenes språkbruk, for å beskrive studentenes muntlige og skriftlige produksjon og for å måle språkutviklingen. (Housen og Kuiken 2009: 461)

Kompleksiteten defineres som «hvorvidt språket som er produsert for å utføre en oppgave er rikholdig og variert»¹ (Ellis 2003: 340 via Housen og Kuiken 2009: 461). Nøyaktighet ses som «evnen til å produsere feilfri tale» og flyt «evnen til å prosessere andrespråket med ‘morsmåls-lignende’ hurtighet» (Lennon 1990: 390 via Housen og Kuiken 2009: 461) eller «hvorvidt språket som produseres for å utføre en oppgave kjennetegnes av pauser, nøling eller omformulering» (Ellis 2003: 342 via Housen og Kuiken 2009: 461). For mer enn 20 år har disse 3 konsepter vært sentrale i andrespråksforskningen og anvendt lingvistikk, men også innenfor kognitiv psykologi og psykolingvistikk, således brukes de ikke bare for å analysere språkbruket og språkutviklingen, men også mekanismer og prosesser som står bakom det språklige uttrykket, språktilegnelsen og bearbeidelsen av kunnskaper om andrespråket (Housen og Kuiken 2009: 462).

På 2000-tallet har diskusjonene rundt CAF speilet en slags krise i forskningsområdet. På den ene siden snakkes det om at vi trenger å definere det vi egentlig vil måle og det som våre forskningsmetoder egentlig viser bedre og mer gjennomtenkt (Pallotti 2009: 599). På den andre siden kalles det for en mer dynamisk og organisk behandling av faget – Larsen-Freeman for eksempel mener at man har hittil fokusert på de enkelte områder av CAF –

¹ Her og heretter min oversettelse.

kompleksitet, nøyaktighet og flyt – altfor separert. Jo nøyere man går inn i disse feltene, jo mer detaljert blir bildet, men også mer komplisert, slik at temaene og spørsmålene tar aldri slutt. Samtidig som vi ser på disse dimensjoner bare én per gang kan vi lett overse at vi har å gjøre med en dynamisk system der delene forandrer seg under tiden, står i samhandling og blir påvirket av hverandre og miljøet rundt dem og i tillegg forandrer denne interaksjonen seg under tiden også. Dermed synes Larsen-Freeman at CAF-forskningen behøver i fremtiden utvikle en bredere teoretisk rammeverk som skulle ta hensyn til at tilegnelsen ikke alltid skjer lineært eller at de tre dimensjoner innenfor CAF er gjensidig avhengige av hverandre og konteksten rundt dem og at de står i dynamisk interaksjon ikke i en statisk tilværelse. (Larsen-Freeman 2009: 582-583, 587)

2.2 Syntaktisk kompleksitet

Ordet «kompleks» eller «kompleksitet» kan ha forskjellige betydninger. På strukturell plan kan vi si at noe er kompleks når den er sammensatt av to eller flere deler eller så kan kompleksiteten peke på varieteten eller at det finnes flere alternativer av noe. Kompleksiteten kan også bety at noe er vanskelig at adskille, analysere eller løse. I språkvitenskapen innebærer dette oftest at betydningen av «kompleks» blir lik med «vanskelig» eller kognitivt sett «krevende». Noen av forskere likestiller kompleksiteten med det som tilegnes senere eller som betraktes som «mer avansert språkbruk». (Pallotti 2009: 593)

Forskjellige forskere har definert kompleksitet på forskjellig måte og med dette følger det helt forskjellige forskningsmetoder og resultat. Pallotti gir en god oversikt (2009: 592-594) over de forskjellige synsvinkler til hva som betraktes som kompleksitet innenfor CAF-forskningen. Først kan kompleksiteten deles i to større kategorier: kompleksiteten av de forskjellige oppgaver (*task related complexity*) som språket brukes til (med dette mener man ikke bare «lekser» i den formelle undervisningen men også bredere kommunikative formål som man vil nå med en viss språklig handling) og kompleksiteten av språket som uttrykkes (*language performance*). Siden min avhandling ikke skal analysere oppgaverelatert kompleksitet, skal jeg ikke gå nærmere inn på det.

Kompleksiteten som beskrives ved selve språklige uttrykket, kan i sin tur betraktes ut fra forskjellige aspekter: leksikalsk, interaksjonell, proposisjonell og forskjellige typer av grammatisk kompleksitet. Samtidig finnes det ikke enighet blant forskere om disse

aspekter bare representerer forskjellige sider av en og samme fenomen eller om de skulle alle sammen behandles som forskjellige konstruksjoner. For eksempel er det problematisk om vi kan i det hele tatt se på leksikalsk og syntaktisk kompleksitet adskilt, siden tilegnelsen av begge to er så tett sammenvevd. Samtidig er ingen av disse underkategorier av kompleksitet enhetlige i seg selv – mange av dem (og blant dem også syntaktisk kompleksitet) er multidimensjonale konsepter med flere underområder. (Pallotti 2009: 593)

I min avhandling har jeg valgt å bruke Lourdes Ortega sin definisjon av syntaktisk kompleksitet (2003: 492):

«Syntactic complexity (also called syntactic maturity or linguistic complexity) refers to the range of forms that surface in language production and the degree of sophistication of such forms.»

Altså syntaktisk kompleksitet betraktes her som både variasjon av forskjellige syntaktiske strukturer og hvor innviklede disse strukturer er. Ortega (2003: 492) legger til at bruk av et slikt begrep i andrespråksforskningen bygger på antakelsen om at andrespråkslæring også innebærer utvikling av det syntaktiske repertoaret og evnen til å bruke dette repertoaret hensiktsmessig i mange ulike situasjoner. For å karakterisere denne utviklingen har man dannet en rekke kvantitative målestokk som gir oss oversikt over lengden av produksjonsenheter, mengden av underordning, omfang av strukturelle typer og hvor avansert de særskilte strukturene er utplassert i produksjonen. Disse målestokk skal presenteres i kapittel 2.4.

Den syntaktiske utviklingen følger vanligvis en viss logisk rekkefølge. Norris og Ortega (2009: 562-263) har beskrevet utviklingsstegene som følgende:

1. I de lavere nivåene av språktilegnelsen uttrykker man seg først og fremst via paratakse / koordinering - sideordning av ord, fraser og setninger.
2. Deretter legger man til hypotakse / underordning for å uttrykke logiske forbindelser mellom ideer.
3. Til slutt begynner man å bruke grammatiske metaforer (i form av forskjellige prosesser som nominalisering eller andre ordklasse-transformasjoner), som fører til avansert språk der man møter mindre underordning og grammatisk kompleksitet,

men som er leksikalsk sett tettere. Slik syntaktisk nivå karakteriseres av mer komplekse setninger i motsetning til flere setninger i ytringen.

Dette eksemplifiseres (eksemplet fra Halliday og Martin 1993/1996 via Norris og Ortega 2009: 563) med følgende sammenlikning av to ekvivalente muligheter for å formidle samme innhold - en 7-ords kompleks T-enhet (1.) og en 6-ords frase (2.):

1. *Darwin thought that species gradually became more complex.*
2. *Darwin's gradual rise to mounting complexity . . .*

Her ser vi hvordan på en høyere nivå av språkbeherskelse kommer det en helt annen type kompleksitet til syne. For å uttrykke samme innhold bruker det første eksemplet underordning, samtidig som det andre bruker nominalisering og erstatning av grammatiske klasser (*became* - verb → *rise* - substantiv; *complex* - adjektiv → *complexity* - substantiv; *gradually* - adverb → *mounting* - adjektiv). Dette er noe som brukes hos de mer avanserte språkbrukere og særlig i skriftlig og akademisk stil. På den måten kan man se at når man uttrykker tettere og lengre ideer i en slik konsentrert form, så blir ytringer faktisk kortere. Den andre frasen kan for eksempel utvides til «*Darwin's gradual rise to mounting complexity is a remarkable intellectual achievement*». (Norris og Ortega 2009: 563) og som vi ser har vi her å gjøre med en enkelytring som ikke inneholder hverken sideordning eller underordning.

Ortega vil framheve at syntaktisk kompleksitet ikke kan brukes som en absolutt målestokk for språkutviklingen, men bare som et utgangspunkt for videre analyser. «Mer kompleks» betyr ikke alltid «mer utviklet» eller «bedre», siden språkutviklingen omfatter mye mer enn bare den syntaktiske utviklingen; man skal ikke heller glemme utviklingen av diskurs og det sosiolingvistiske repertoaret. Dermed er syntaktisk komplekse tekster ikke alltid «gode» tekster og omvendt «gode» tekster bør ikke alltid være syntaktisk komplekse. (Ortega 2003: 494)

Undersøkelser om andrespråksskrivning der syntaktisk kompleksitet står sentralt har blitt brukt for å evaluere virkningene av pedagogiske intervensjoner på utviklingen av grammatikk eller skriveferdigheter, for å undersøke oppgave-relatert variasjon i andrespråksskrivning og vurdere forskjeller i tekster skrevet på tvers av ferdighetsnivåer og over tid (Ortega 2003: 492-493). I denne avhandlingen har disse målestokk blitt brukt

for å analysere andrespråkstekster på to forskjellige utviklingsnivåer, samt sammenlikne dem med førstespråkstekster.

2.3 Forklaringer for noe sentrale begrep

I min analyse bruker jeg forskjellige syntaktiske «størrelser» når jeg diskuterer syntaktisk kompleksitet. Siden forskjellige undersøkelser har brukt terminologien litt variert, finner jeg det viktig å definere hovedbegrepene nærmere, gi deres engelskspråklige ekvivalenter og forklare hvordan jeg bruker dem i analysene.

Ytring (*Orthographic sentence / Utterance*) og enkelytring (*Simple sentence*)

Den største syntaktiske enheten som jeg bruker i min analyse, er ytring. En ytring er en syntaktisk enhet som begynner med storbokstav og ligger som vanlig mellom store skille tegn (Faarlund et al 1997: 923). Jeg har i tillegg valgt å bruke begrepet enkelytring (*simple sentence*) for ytringer som bare består av en setning, dvs. som ikke inneholder leddsetninger eller sideordnede setninger.

For eksempel «*Han sa at norsk syntaks ikke var så vanskelig å lære seg og at den egentlig likner mye med estisk.*» er en ytring, samtidig som «*Norsk syntaks er lett å lære seg.*» er en enkel ytring.

T-enhet (*T-unit / Minimal terminal unit*)

T-enhet er et begrep foreslått av Kellogg Hunt på 1960-tallet for kartlegging av barnas utvikling av syntaktisk kompleksitet på førstespråket. Begrepet har etter det blitt flittig brukt både i første- og andrespråksforskning: både ved analyse av muntlig (Bardovi-Harlig 1992: 390) og skriftlig språkbruk (Larsen-Freeman 1978: 441, Gaies 1980: 53). Jeg har valgt å bruke målestokk som antall T-enheter i teksten og lengden av T-enheter slik at mine resultat vil bli sammenlignbare med andre liknende undersøkelser utført i feltet både hittil og i fremtiden.

Hunt definerte T-enheten som den korteste enheten som et stykke diskurs kan deles i, uten at noen setningsfragmenter skulle bli etterlatt (Jiang 2013: 2). Det vil si at en T-enhet består av bare en selvstendig setning (hovedsetning) og alle underordnede leddsetninger som er knyttet til den (i fall de finnes), mens sidestilte setninger utgjør hver sin T-enhet.

For eksempel «*Jeg er en student ved Skandinavistika og jeg har lært norsk i 5 år.*» består av 2 T-enheter (6 ord + 8 ord), mens «*Min venn som studerer økonomi har gått på norsk kveldskurs i 3 år.*» består av 1 T-enhet (13 ord). Om vi bare skulle se på antallet ord i ytringene så viser det seg at begge ytringer er nesten like lange – hhv. 14 og 13 ord. Således sier ytringslengde ikke mye om den syntaktiske kompleksiteten. Samtidig om vi regner ut antallet ord i T-enheter, kan vi se at lengre T-enheter pleier å være mer komplekse og inneholde mer underordning.

Setning (Clause)

En setning er den minste analyseenheten i avhandlingen (lavere nivåer som strukturer innenfor setningsledd, fraser og setningsfragmenter har jeg valgt til å se bort fra i den kvantitative analysen). Norsk Referansegrammatikk definerer setningen som en forbindelse av et nominal (setningens subjekt) og en verbfrase (et predikat med en finitt verb som kjerne) (Faarlund et al 1997: 39). Setninger kan fungere som hovedsetninger, sidestilte setninger eller leddsetninger. Ved analyse av tekstene har jeg delt tekstene i setninger ved å gå ut fra den enkle reglen om ett finitt verb – én setning, dvs. at en setning bare inneholder ett finitt verb.

For eksempel ytringen «*Min mormor som bor i Syd-Estland fyller 90 år og vi planlegger å markere det med en stor fest.*» består av 2 t-enheter, men 3 setninger – 2 sidestilte setninger og en leddsetning. Dermed kan vi se at når vi ser bare på antall setninger per ytring, så viser det at ytringen er satt sammen av enten underordnede eller sideordnede setninger. Ved å sammenlikne antallet setninger i ytringen med antallet T-enheter i ytringen, så ser vi at jo mer setninger og jo mindre T-enheter det finnes per en ytring, jo mer underordnede setninger det må være i teksten.

Hovedsetning og oversetning (Main clause)

Hovedsetning er en fullstendig grammatisk enhet som har oppfylt kravene til en setning og som kan stå alene mellom to store skilletegn. Hovedsetningen kan inneholde en eller flere leddsetninger. Oversetning er setningen som leddsetningen er en del av, det kan både være en hovedsetning eller en annen leddsetning. (Faarlund et al 1997: 41)

For eksempel «*Jeg syns at du bør spise mer grønnsaker fordi det er sunt.*» er en hovedsetning som inneholder 2 leddsetninger. *Jeg syns* er oversetningen til *at du bør spise mer grønnsaker* og dette i sin del er oversetningen til *fordi det er sunt*.

Den tradisjonelle definisjonen av en hovedsetning omfatter både enkelytringer, sideordnede setninger og hovedsetninger som inneholder leddsetninger. I den kvantitative analysen har jeg begrenset definisjonen av en hovedsetning til bare den setningen som har enten en sideordnet setning eller leddsetning knyttet til seg (i eksempelet ovenfor er det da bare *Jeg syns*). Dermed har jeg ikke markert enkelytringer eller sideordnede setninger under kategorien hovedsetning (selv om de selvfølgelig bruker hovedsetnings leddstilling – se nedenfor). En slik begrensning er nødvendig for å spesifisere hvor de forskjellige syntaktiske strukturer eksakt forekommer og se om det, at man har lagt til en sideordnet eller underordnet konstruksjon til en setning, har påvirket setningsstrukturen eller den syntaktiske korrektheten i hovedsetningen. Enkelytringer og sideordnede setninger har jeg markert og analysert separat.

Leddsetning (*subordinate clause*)

Leddsetning er en setning som ikke kan stå alene og som fungerer som ledd i en annen setning. (Faarlund et al 1997: 41) En hovedsetning kan inneholde flere leddsetninger som er underordnet hverandre (første-, andre-, tredjegradsunderordning osv). Leddsetninger kan ha forskjellige funksjoner innenfor oversetningen. Ut fra hvilket ledd de utgjør i oversetningen har man kategorisert leddsetninger i forskjellige kategorier: nominale, adverbiale og adjektiviske leddsetninger (Golden 2011: 145, 148, 151). Leddsetningen kan stå i begynnelsen av en oversetning (foranstilt leddsetning), midt i en oversetning (innskudd leddsetning) eller ved slutten av en oversetning.

Nominale leddsetninger er leddsetninger som fungerer som nominale ledd i oversetningen. Vanligvis fungerer nominale leddsetninger som objekt, men de kan stå for andre nominale ledd som subjekt eller predikativ også. Følgende hovedsetning inneholder en nominal leddsetning som fungerer som objekt: «*De fortalte at hun hadde vunnet*»; leddsetningen som fungerer som subjekt finner vi her: «*Hva som har skjedd er uklart.*» (Golden 2011: 145)

Adverbiale leddsetninger fungerer i oversetningen som et adverbial. For eksempel her: «*De fortalte det da de kom*». (Golden 2011: 148) Adverbiale leddsetninger kan deles inn i underkategorier etter betydning (for eksempel tidssetninger, årsakssetninger, motsetningssetninger osv). Kategoriene speiler mer det semantiske innholdet i leddsetningen, samtidig som setningsstrukturen og den syntaktiske kompleksiteten ikke er

avhengig av hvilken type adverbial leddsetning vi har å gjøre med. Derfor har jeg ikke inkludert disse underkategorier i min analyse.

Adjektiviske leddsetninger er leddsetninger som ikke utgjør et selvstendig ledd i oversetningen, men som er knyttet til et nominalt ledd i oversetningen. Allikevel betraktes dem som leddsetninger eftersom leddstillingen følger leddsetningsskjemaet. Adjektiviske leddsetninger kan utgjøre en del av forskjellige nominale setningsledd. De kan være:

- del av subjektet: *Bilen som vi vant, er blitt stjålet.* Her står *som* som subjekt i leddsetningen og den adjektiviske leddsetningen beskriver subjektet i oversetningen;
- del av objektet: *Vi traff en mann (som) jeg kjenner.* Her står *som* som objekt i leddsetningen og den adjektiviske leddsetningen beskriver objektet i oversetningen;

Adjektiviske leddsetninger kan også stå til et ord for tid eller sted:

- presisering av tid: *Jeg husker godt den sommeren da vi dro på seilferie.* Man må legge merke til at her utgjør leddsetningen ikke et adverbial i oversetningen, men presiserer objektet *den sommeren*.
- presisering av sted: *Det huset der de bodde, er nå ødelagt. Du kan sitte her hvor det er varmt.* I den første ytringen presiseres det subjektet i oversetningen, i den andre ytringen presiseres det adverbialet.

Subjunksjonen (for eksempel *som*) i den adjektiviske leddsetningen kan sløyfes i fall den ikke fungerer som subjekt i setningen.

Koordinering, sideordnet setning (compound sentence)

Koordinering eller sideordning binder sammen setningsleddene eller setninger vanligvis ved hjelp av konjunksjoner. Både hovedsetninger og leddsetninger kan sideordnes, men sideordning skjer alltid på samme nivå. Sideordning kan skje innenfor en ytring eller mellom to separate ytringer. Jeg har valgt å holde meg innenfor ytringsgrenser, således dekker min analyse bare sideordning av setninger i en og samme ytring. For eksempel: «*Han var trøtt, men han ville ikke legge seg ennå.*» er en ytring som består av 2 setninger sideordnet med konjunksjonen *men*. (Golden 2011: 108-109)

Inversjon

Inversjon oppstår når man setter et annet ledd enn subjektet i Forfeltet (se side 14

«Ordstilling i setninger» nedenfor). Da kan man ikke bruke SVO rekkefølgen som er vanlig i norsk, siden V2-regelen (finit verb på andre plass i setningen) medfører at subjektet skal settes bak verbet (Hagen 1998: 272; Johansen 2008: 50). Inversjon kan man bare bruke i hovedsetningen, siden leddsetningsrekkefølgen er fast og vi kan ikke ha noe annet enn subjektet på førsteplass. For eksempel: «Jeg har ikke fått noe brev i dag.» blir invertert når vi setter adverbialet først i setningen «I dag har jeg ikke fått noe brev.», men dette kan ikke gjøres i leddsetningen. Vi må ha subjekt på førsteplass i leddsetningen «Jeg så at jeg ikke hadde fått noe brev i dag.», men vi kan ikke plassere adverbialet først «Jeg så at i dag hadde jeg ikke fått noe brev. *»

Inversjon har vært i sentrum i andrespråksforskningen fordi det er et trekk som synes å være vanskelig å tilegne seg og det regnes ofte som en komplisert syntaktisk struktur (Johansen 2008: 50). Derfor har jeg i tillegg til å analysere forskjellige typer av setninger, også valgt å markere separat T-enheter som inneholder inversjon og analysere om studentene har klart å invertere korrekt. Inversjon som er blitt brukt i en leddsetning har jeg markert som mislykket. Det er viktig å forstå at setninger som inneholder inversjon ikke er noen separat setningstype i seg selv. Inversjon er heller en ekstradimensjon av syntaktisk kompleksitet som kan forekomme ved andre mer eller mindre komplekse strukturer (for eksempel i enkelytringer, i hovedsetninger og i sideordnede setninger). Ofte møter man inversjon også i forbindelse med en foranstilt leddsetning som for eksempel: «Når jeg kom hjem, så jeg at jeg ikke hadde fått noe brev.»

Setninger med formelt subjekt

Siden norsk har både verb- og subjektskrav i setningen (jf. definisjonen ovenfor), må det finnes et subjekt i hver setning. Da vi hverken kan eller vil ikke uttrykke hvem eller hva som foretar handlingen, må vi bruke et formelt subjekt – *det*. Formelt subjekt brukes også ved såkalte presenteringskonstruksjoner, i så fall får vi to subjekter i setningen – formelt subjekt *det* og det egentlige subjektet som alltid står i ubestemt form. (Golden 2011: 133, 157) For eksempel i setningen «Det ligger en ny bok på bordet» er *en ny bok* det egentlige subjektet. På samme måte som ved inversjon, må det egentlige subjektet plasseres etter verbalet, den gangen i slutfeltets nominal-plass. Forskjellen mellom setninger med formelt subjekt og setninger med inversjon er at i den førstnevnte plasserer vi i forfeltet ikke noe som vi har hatt i setningen før, men noe (det innholdstomme *det*) som bare må settes inn for å fylle i forfeltet. Jeg har dermed valgt å betrakte slike setninger som

komplekse strukturer på lik linje med setninger som inneholder inversjon. På samme måte som setninger med inversjon ikke er et separat setningstype, så er formelt subjekt også noe som kan forekomme i alle de andre setningstypene og som snarere danner et ekstradimensjon av setningskompleksitet.

Tabellen nedenfor illustrerer sammenfatningsvis forholdene mellom de syntaktiske enhetene som jeg har brukt i min analyse. Eksempelet består av en hovedsetning som har en foranstilt leddsetning og en sideordnet setning knyttet til seg. I hovedsetningen har vi i tillegg inversjon og et formelt subjekt. Ytringen inneholder 2 T-enheter og 3 setninger.

Tabell 1: *Forholdene mellom syntaktiske enheter*

<i>Når du lær deg norsk, kan det være vanskelig i begynnelsen, men etterpå blir det enklere.</i>		
Leddsetning	Hovedsetning	Sideordnet setning
T-enhet		T-enhet
Ytring		

Ordstilling i setninger

Norsk har forholdsvis fast ordstilling, dvs. setningsleddene kommer alltid i en viss rekkefølge og ved syntaktiske analyser bruker man setningsskjemaer for å bestemme om setningen er syntaktisk korrekt eller ikke. Leddstillingen i hovedsetninger og leddsetninger er forskjellig. Hovedsetningsskjemaet ser ut slikt:

Tabell 2: *Skjema for hovedsetninger (Golden 2011: 163)*

Forfelt	Midtfelt				Slutfelt		
F	v	a1	n	a2	V	N	A
	Verbal	Adverbial	Nominal	Adverbial	Verbal	Nominal	Adverbial
Forskjellige setningsledd. Finnes bare i hovedsetninger.	Den finitte delen av verbalet	Setningsadverbial	Subjekt, trykklette pronomen (eventuelt som objekt)	Setningsadverbial	Den infinitte delen av verbalet	Direkte og indirekte objekt, predikativ	Måte, sted, tid
<i>Jeg</i>	<i>har</i>	<i>ikke</i>			<i>fått</i>	<i>noe brev</i>	<i>i dag.</i>
<i>Sannsynligvis</i>	<i>skal</i>		<i>de</i>		<i>bestemme</i>	<i>hvem som vant konkuransen</i>	<i>i morgen.</i>

Leddsetninger har et fast setningsstruktur på norsk. De innledes av en subjunksjon og de har ikke noen forfelt der vi kan legge forskjellige setningsledd som i hovedsetningen, leddene i midtfeltet har også litt annerledes rekkefølge. Leddsetningsskjemaet ser ut slikt:

Tabell 3: Skjema for leddsetninger (Golden 2011: 165)

Forbinderfelt	Midtfelt			Sluttfelt		
Fb	n	a	v	V	N	A
Subjunksjon	Nominal	Adverbial	Verbal	Verbal	Nominal	Adverbial
	Subjekt	Setnings- adverbial	Finitt del	Infinitt del	Direkte og indirekte objekt, predikativ	Måte, sted, tid
<i>at</i>	<i>jeg</i>	<i>ikke</i>	<i>har</i>	<i>fått</i>	<i>noe brev</i>	<i>i dag.</i>
<i>siden</i>	<i>de</i>	<i>sannsynligvis</i>	<i>skal</i>	<i>bestemme</i>	<i>hvem som vant konku- ranssen</i>	<i>i morgen.</i>

2.4 Målestokk for syntaktisk kompleksitet

Materialet har blitt analysert både med hjelp av kvantitative og kvalitative metoder. Først har den syntaktiske kompleksiteten blitt kartlagt og deretter beskrevet med hjelp av kvantitative målestokk. Basert på funnene fra den kvantitative analysen, har jeg analysert nærmere et utvalg eksempler av trekk der vi kan se forskjeller i produksjon mellom de to studentgrupper eller mellom studentene og kontrollgruppen. Mulige forklaringer for forskjeller og problemområder har også blitt presentert.

For den kvantitative analysen har jeg først delt ytringer i tekstene inn i setninger og T-enheter, talt og regnet ut lengden av ytringer, T-enheter og setninger og markert setningene i henhold til forskjellige syntaktiske kategorier (sideordning, underordning, forskjellige typer av leddsetninger osv). Målestokk som jeg har brukt er blitt brukt i andre undersøkelser innenfor andrespråksforskning (Ortega 2003), dvs resultatene (gjennomsnittsverdier, andel i prosent) vil bli sammenlignbare med tidligere og framtidige forskningsresultater både innenfor norsk som andrespråk men også andre språk.

Samtidig som disse målestokk har blitt brukt på lenge, har det også i det siste vært diskusjon rundt det hvordan kompleksiteten og særlig den syntaktiske kompleksiteten skal måles og om forskningsdesignene har tilfredsstillende og tilstrekkelig klart å beskrive kompleksiteten. Ortega og Norris (2009: 558) oppfordrer oss å tenke på hva vi egentlig prøver å måle og hvorfor. De viser at de mest brukte målestokk for syntaktisk kompleksitet handler om lengden av de forskjellige syntaktiske enheter og mengden av underordning. Det er bare noen få som også tar hensyn til mengden av sideordning og noen som har prøvd å kombinere disse målestokk med å se på frekvensen, varieteten og tilegnelsestidspunktet av visse grammatiske former (for eksempel passive eller infinitte konstruksjoner, kondisjonalis osv.) (Ortega og Norris 2009: 559). Samtidig påstår de at noen av målestokkene som brukes, egentlig måler et og samme fenomen, samtidig som noen andre fenomen ofte er ekskludert fra analysene. For eksempel både ‘andelen av underordnede setninger per alle setninger sammenlagt’ og ‘andelen av setninger per en T-enhet’ viser underordningsgraden, og således dublerer de hverandre. Samtidig utelates det andre viktige momenter ved den syntaktiske kompleksiteten. Når vi skulle analysere språkutviklingen bare ut fra disse mål, kunne vi se en økning i måleverdiene bare i fall det blir produsert mer underordnede setninger. Ortega og Norris mener at på den måten skulle vi se syntaktisk kompleksitet (som er egentlig en multi-dimensjonell fenomen) ekstremt snevert og miste eller i verste fall feilinterpretare noen viktige momenter ved språktilegnelsen. (2009: 560, 566).

Av den grunn mener Norris og Ortega at det er viktig at CAF-forskningen også skal omfatte forskjellige dimensjoner av syntaktisk kompleksitet. Hva som skaper syntaktisk kompleksitet har de delt i fem underkonstruksjoner - alle bør tas til hensyn ved alle undersøkelser for å forsikre sammenlikningsgrunnlaget, samtidig er noen av disse som er særlig relevante ved visse språkutviklingsnivåer (Norris og Ortega 2009: 561-562):

- **generisk kompleksitet** som måles ved å se på lengden av produksjonsenheter og som kan brukes ved alle tilegnelsesnivåer;
- **kompleksitet via sideordning** som er spesielt relevant ved undersøkning av lavere språkbeherskelsesnivåer;
- **kompleksitet via underordning** som passer best ved undersøkning av mellomnivået;
- **setningsintern kompleksitet** som uttrykker seg først ved mer avanserte språkbrukere;

- **kompleksitet sett som strukturell variasjon og raffinement**, dvs. hvor innviklede former og hvor variert man bruker og når de forekommer – dette er igjen noe som kan betraktes gjennom hele utviklingsløpet.

Norris og Ortega (2009: 564, 574) hevder at for å få tilfredsstillende resultat som kan sammenliknes på tvers av alle utviklingsnivåer, er minimumkravet til undersøkelsen at det skulle måle både sideordning, underordning og setningsintern kompleksitet. Samtidig mener de at det også er viktig å beskrive andre sider av syntaktisk kompleksitet for å få en helhetlig forståelse av fenomenet. Dette har jeg prøvd å ta hensyn til ved valg av mine analysekategorier.

For å beskrive de forskjellige aspektene ved syntaktisk kompleksitet, trenger man å måle dem på noen måte. Sentralt i min analyse står det 6 kvantitative målestokk, som er blitt flittig brukt i andrespråksforskningen. Ortega har gjennomført en syntesestudie av 21 studier om syntaktisk kompleksitet og baserer sin analyse på disse 6 mest brukte mål på tvers av studiene (2003: 496):

- **gjennomsnittslengden av ytringer** (*MLS – mean length of sentence*),
- **gjennomsnittslengden av T-enheter** (*MLTU – mean length of T-unit*)
- **gjennomsnittslengden av setninger** (*MLC – mean length of clause*)
- **gjennomsnittsandelen av T-enheter per ytring** (*TU/S – T-unit per sentence*)
- **gjennomsnittsandelen av setninger per T-enhet** (*C/TU – clauses per T-unit*)
- **gjennomsnittsandelen av leddsetninger per setninger** (*DC/C – dependent clauses per clause*)

Målene som gjennomsnittslengden av ytringer, gjennomsnittslengden av T-enheter og gjennomsnittslengden av setninger gir oss oversikt over lengden av produksjon på setnings og frasenivå. Det er selvklaart at bare ved å se på lengden av ytringer og T-enheter gir oss en begrenset bilde av syntaktisk kompleksitet. Begge to kan gjøres lengre ved å legge til leddsetninger, men ytringer kan også bli lengre når de inneholder mer koordinasjon og begge enheter kan forlenges ved å legge til adjektiver, preposisjonsfraser, infinitte konstruksjoner eller pre- eller postmodifikasjoner av substantiver og så videre. Derfor kan lengden av ytringer og T-enheter brukes bare som en global eller generisk målestokk for syntaktisk kompleksitet. (Norris og Ortega 2009: 561) Det vil si at lengden av ytringer og T-enheter sikkert formidler oss noe om den syntaktiske kompleksiteten (korte ytringer

pleier å være mindre komplekse), men uten å se nærmere på det, kan vi ikke påstå hvor den syntaktiske kompleksiteten (eller enkelheten) ligger. Lengden av setninger er derimot en mer spesifikk mål som viser oss kompleksitet på et mindre nivå - frasenivå, eftersom setninger kan bli lengre bare ved å utvide dem med adjektiver, adverbier, preposisjonsfraser, infinitte konstruksjoner eller gjennom nominalisering.

Samtidig som økning i disse 3 målestokk viser oss økt kompleksitet, må man være litt forsiktig med sine konklusjoner. Som nevnt før: «lengre» betyr ikke alltid «mer kompleks» og «mer kompleks» betyr ikke alltid «bedre» eller «mer utviklet». Flere undersøkelser har vist at økning i underordningsgraden kan betraktes bare frem til en viss tidspunkt i språktilegnelsen, etter dette stabiliserer det seg, samtidig som lengden av T-enheter fortsetter å øke (dvs. setningsintern kompleksitet øker). På et avansert nivå ser man allerede mindre underordning enn i tekstene på mellomnivå (Norris og Ortega 2009: 564) og dette kan skyldes generellt (ikke bare syntaktisk) økt kompleksitet i språkproduksjonen (Norris og Ortega 2009: 566). For å vurdere setningsintern kompleksitet, har jeg målt lengden av setninger, i tillegg har jeg utført en kvalitativ analyse av et utvalg eksempelsetninger.

Underordningsgraden er allikevel et viktig del av den syntaktiske kompleksiteten. For å måle den har jeg brukt 2 målestokk: gjennomsnittsandelen av setninger per T-enhet og gjennomsnittsandelen av leddsetninger per setninger som begge beskriver mengden av underordning (Ortega 2003: 496). Jo større disse mål er jo mer underordnede konstruksjoner har blitt brukt i teksten.

Noe som ikke har fått mye oppmerksomhet i andrespråksforskningen er kompleksiteten som skapes via sideordning. Dette er allikevel særlig relevant for å se på når man analyserer lavere nivåer av språktilegnelsen. (Norris og Ortega 2009: 562) En av målestokkene som kan brukes for å gjenspeile sideordningsgraden er gjennomsnittsandelen av T-enheter per ytring (Ortega 2003: 496). Hvis en ytring består av flere T-enheter, må de være sidestilte og oftest også sammenbundet med konjunksjoner. Det betyr at jo større dette mål er, jo mer koordinasjon brukes i teksten. I min avhandling betrakter jeg sideordnede setninger som komplekse strukturer, selv om vi kan si at de på norsk ikke er like komplekse som leddsetninger (på grunn av restriksjoner og det forskjellige leddstillingen i leddsetninger). Allikevel er både side- og underordnede setninger mer

komplekse og krever med kognitiv prosessering enn enkelytringer. Et eksempel på hvordan kompleksiteten øker seg gjennom de forskjellige typer setninger:

1. *Det regner. Jeg kan ikke gå hjem.* (2 enkelytringer)
2. *Det regner, så jeg kan ikke gå hjem.* (Hovedsetning + sideordnet setning)
3. *Jeg kan ikke gå hjem, ettersom det regner.* (Hovedsetning + leddsetning)
4. *Siden det regner, kan jeg ikke gå hjem.* (Foranstilt leddsetning + hovedsetning med inversjon)

For å ta hensyn til Norris og Ortega sitt kritikk mot det at det er veldig få undersøkelser som ser på varieteten av de forskjellige former som brukes for å skape syntaktisk kompleksitet (2009: 562), har jeg utvidet min analyse til et mer detaljert nivå. For å få et bedre forståelse av hva slags strukturer skaper syntaktisk kompleksitet i tekstene, har jeg markert forskjellige typer av komplekse strukturer separat. Det vil si at jeg har ikke bare analysert i hvilken grad komplekse strukturer er blitt brukt, men også hvilke typer av komplekse strukturer studentene bruker og hvor mye de bruker dem. Analysekategoriene inneholder forskjellige typer leddsetninger, men også inversjon og formelt subjekt. Valg av kategoriene går direkte ut fra det som faktisk finnes i de konkrete tekstene. Jeg har inkludert bare kategoriene / fenomener som tekstene inneholder og la til eller modifiserte kategoriene underveis når jeg fant noe nytt i teksten.

Målestokk som viser variasjon og utbredelse av forskjellige komplekse syntaktiske strukturer:

- Andelen av nominale leddsetninger per leddsetninger.
 - o Andelen av nominale leddsetninger som objekt.
 - o Andelen av nominale leddsetninger som subjekt.
- Andelen av adjektiviske leddsetninger per leddsetninger
 - o Andelen av adjektiviske leddsetninger der *som* fungerer som objekt
 - o Andelen av adjektiviske leddsetninger der *som* fungerer som subjekt
 - o Andelen av adjektiviske leddsetninger som tilhører en tidsuttrykk
 - o Andelen av adjektiviske leddsetninger som tilhører en stedsuttrykk
- Andelen av adverbiale leddsetninger per leddsetninger
- Andelen av T-enheter med inversjon per alle T-enheter
- Andelen av setninger med formelt subjekt

Til slutt har jeg sett på hvor framgangsrikt studentene danner syntaktisk komplekse konstruksjoner og i hvilke syntaktiske kategorier man har mest utfordringer. Ved analysen markerte jeg for enhver setning om setninger var syntaktisk korrekt eller ikke. Hva som er normen til norsk syntaks, har jeg basert på Golden *et al.* (2011) «Norsk som fremmedspråk. Grammatikk.» og «Norsk referansegrammatikk» (Faarlund *et al.* 1997), samt konsultert med morsmålstalende spesialister.

Siden fokuset i avhandlingen er satt på den syntaktiske kompleksiteten, har jeg analysert kun det om setningen var *syntaktisk* korrekt eller ikke, og dermed har jeg valgt og markere setninger som korrekte også i fall studenten har avviket fra målspråksnormen på grunn av manglende kunnskaper innenfor morfologi eller ordforråd (for eksempel manglende preposisjoner, feil valg av tempus eller verbets valens, bestemthet osv.), dette selvfølgelig dersom setningsstrukturen og setningsleddenes plassering var korrekte. Her er for eksempel en ytring fra gruppe B: *Jens var litt sint, men samtidig også lykkelig at det ikke skjedde noe vondt med hunden.* Vi har å gjøre med en syntaktisk kompleks ytring, som består av en sideordnet og en underordnet setning. Studenten har litt leksikalske problemer som gjør at det mangler en preposisjon i setningen (det skal stå enten *glad for at* eller *lykkelig over at*) og valg av tempus (skulle brukes preteritum perfektum *at det ikke hadde skjedd noe...*). Til tross for det er ytringen syntaktisk sett helt korrekt, setningsleddene står i rett rekkefølge.

Bestemmelser for hvor grensen mellom syntaktisk korrekte og syntaktisk feilaktige setninger går, var ikke alltid lett, særlig i nybegynnergruppen, der setningene kan vise problemer i veldig mange språkområdene samtidig. I de vanskelige tilfeller har jeg allikevel prøvd å identifisere setningsleddene og se på hvilken funksjon ordene eller frasene har i setningen. Ved å konsultere med andre spesialister og støtte på min egen sunn fornuft var det mulig å markere også disse setninger. Likevel må jeg erkjenne at avvikene fra syntaktiske normer er særlig i de lavere tilegnelsesnivåene tett knyttet til at man ikke helt har tilegnet seg alle leksikalske og morfologiske nyanser – dette ville være et interessant utgangspunkt for videre studier. På grunn av masteravhandlingens begrenset omfang har jeg likevel valgt å utelate den diskusjonen.

For å analysere tilegnelsen av de forskjellige syntaktiske strukturer har jeg brukt følgende målestokk:

- Andelen av syntaktisk korrekte t-enheter
- Andelen av syntaktisk korrekte enkelytringer
- Andelen av syntaktisk korrekte hovedsetninger
- Andelen av syntaktisk korrekte sideordnede setninger
- Andelen av syntaktisk korrekte leddsetninger (og deres underkategorier – nominale, adverbiale og adjektiviske leddsetninger)
- Andelen av syntaktisk korrekte setninger med inversjon
- Andelen av syntaktisk korrekte setninger med formelt subjekt *det*.

3 Materialet og informanter

Materialet som analyseres i oppgaven består av skriftlige tekster fra norskstudentene ved Universitetet i Tartu. Tekstene ble samlet inn fra to grupper av voksne studenter. Ved tekstenes innsamlingstidspunktet hadde den ene gruppen (gruppe A) lært seg norsk bare i et halvt år, 2 timer (2×90 min) i uka (selvrapportert gjennomsnitt 6,6 måneder) og de studerer norsk som et valgfritt kurs. Den andre gruppen (gruppe B) studerer norsk som hovedfag og hadde ved innsamlingstidspunktet lært det i 4 semestre ved å delta på forelesninger og seminar i de forskjellige norskspråklige fag (selvrapportert gjennomsnitt 23,2 måneder). Alle informanter er voksne studenter som har norsk som sitt tredje, fjerde, femte eller til og med sjette fremmedspråk.

Ved siden av de 2 gruppene ble det brukt tekster innsamlet fra en mindre «kontrollgruppe» av innfødte norsktalende (gruppe C). Det kan ikke helt betraktes som en kontrollgruppe, siden vi ikke har å gjøre med en eksperimental undersøkelse der vi skulle ha utført noen intervensjoner. Likevel fungerer gruppe C som et sammenliknings og evalueringsgrunnlag for å se hvilken nivå av syntaktisk kompleksitet morsmålsbrukere anvender for å utføre den samme oppgaven. Viktigheten av å bruke data fra innfødte talere av målspråket understrekes også av Pallotti (2009: 598), ikke nødvendigvis fordi målet til språktilegnelsen skulle være å oppføre seg på eksakt samme måte som morsmålsbrukere, men heller for å komme seg bort fra holdninger som synes at studentene er på noen måte mangelfulle språkbrukere som alltid må gjøre noe mer og noe bedre. Diskusjonen rundt dette har jeg fordypet meg i ved teorikapitlet.

 herell oversikt over gruppenes størrelse og tekstenes omfang vises i tabell 4.

Bildeserien har blitt omfattende brukt i lingvistisk forskning blant annet i undersøkelser om språktypologi og kontrastiv lingvistikk (Slobin 2004), førstespråkstilegnelse og barnespråk (Berman og Slobin 1994), språkvansker (Reilly et al 2004), andrespråkstilegnelse (Tortani 2012) og tospråklighet (Bennet-Kastor 2002). Den har blitt mest brukt for å undersøke muntlig produksjon i form av narrative fortellinger og oftest ved undersøkelse av barnespråk, men den er også godt tilegnet for undersøkelse av skriftlig produksjon.

Tabell 4: Generell informasjon om informantgruppene

	Karakteristikk	Hvor lenge har de lært norsk?	Antall informanter	Antall ord	Antall ord per tekst	Antall setninger	Antall T-enheter
Gruppe A	Nybegynnergruppen som har norsk som et valgfritt kurs	6,6 måneder	11	2240	204	454	377
Gruppe B	Andreårsstudenter med norsk som hovedfag	23,2 måneder = 4 semester	10	4024	402	732	524
Gruppe C	Innfødte norsktalere	Morsmål	5	3046	609	480	480

Siden alle tekster har samme bildeserie som grunnlag – det er samme steder, figurer og handlingsgang som må beskrives – er tekstene godt sammenlignbare. Studentene skrev tekstene i sine norsktimene på universitetet, men oppgaven var ikke en del av deres evaluering. De ble informert om at de ikke skal få noen karakter for den og at tekstene brukes for vitenskapelig undersøkelse. Det var dermed ikke en situasjon der studentene skulle bli påvirket av prestasjonsangst, men samtidig var de bevisst på at tekstene skal analyseres og dette kan ha påvirket dem. Det ble ikke satt noen tidsgrenser for skriveprosessen, men i prinsipp hadde studentene 1,5 timer for å bli ferdig med fortellingen. De fleste klarte det fint og de hadde ikke særlig tidspress på seg.

For å forsikre at studentene ikke skulle gå i stå med sin skrivning, ble de forsynet med et utvalg substantiv (oversatt fra estisk til norsk) som var absolutt nødvendige for å kunne fortelle historien (se tabell 5). Siden min analyse bare gjelder syntaksen i tekstene så vil dette ikke påvirke mine funn, snarere omvendt – jeg fikk mer materiale til å analysere fordi studentene kunne beskrive mer av hendelsene på bildene enn de kunne ha ellers gjort med sine begrensede ordforråd.

Tabell 5: Ord som studentene ble utstyrt med på forhånd

konn - en frosk, -er	et hvepsebol
purk - et glass, -	mutt - en muldvarp
mesilane - en bie, -r	öökull - en ugle, -er
en bikube	hirv - en hjort
herilane - en veps, -er	å kvakke

Studentene skrev tekstene for hånd. For å gjøre analysen av tekstene smidigere, ble tekstene deretter digitalisert. Siden håndskriften ikke alltid var forståelig, ble det brukt følgende markeringer:

#gjennomstreking#

[.] forsiktig skrevet / lyst

? uforståelig bokstav

<ikke forståelig>

/lagt til/

Digitaliserte tekstene ble delt inn i setninger og hver setning analysert og kategorisert. Antall ord i de forskjellige måleenheter ble regnet, setningene plassert i analysekategoriene og deretter bestemt om de var syntaktisk sett korrekte eller ikke. Setningsanalysen ble utført i en Excel-tabell. Analysekategoriene ble nærmere presentert i kapittel 2.4.

4 Analyse og resultater

4.1 Nybegynnergruppen (Gruppe A)

4.1.1 Generelt om tekstene

Materialet for gruppe A omfatter 11 tekster. Følgende tabell gir oversikt over de viktigste karakteristika av tekstene og gruppen:

Tabell 6: *Tekstene i Gruppe A*

		Maksimum	Minimum
Antall tekster	11		
Antall ord sammenlagt	2240		
Ord per en tekst	204	318	116
Antall T-enheter sammenlagt	377		
T-enheter per en tekst	34		
Antall setninger sammenlagt	454		
Setninger per tekst	41		
Antall ytringer sammenlagt	373		
Ytringer per tekst	21		

Som det framgår av tabell 6 er tekstene for gruppe A ikke veldig lange (se sammenlikning med andre grupper i kapittel 4.3.1, tabell 10), noe som sikkert speiler deres fortsatt begrensede språkkunnskaper. På bakgrunn av grundig gjennomlesning av alle tekstene kan jeg si at selv om studentene i gruppe A av og til har vært ganske lakoniske i deres beskrivelse av hendelser på bildene, så har alle klart å få den viktigste informasjonen med og formidlet historien tydelig. Sammenlagt skaper materialet en ganske omfattende tekstmasse, som er tilstrekkelig for å analysere de syntaktiske kategoriene og lage generaliseringer. Som vi ser i tabell 6 er forskjeller i antall ord, T-enheter, setninger og ytringer per tekst ganske store mellom den korteste og lengste teksten i gruppen. Samtidig fordeler de andre tekstene seg ganske jevnt mellom disse ytterpunkter. I tabell 6 har jeg gitt gjennomsnittsverdier, men jeg sjekket mine resultat også mot medianverdier og det viste seg ikke å være store forskjeller mellom gjennomsnitts- og medianverdier (for eksempel antall ord per tekst: 204 i gjennomsnitt, 195 i median; antall T-enheter per tekst både gjennomsnitt og median 34).

Analyse og markering av setningene var ikke alltid ukomplisert ved gruppe A. Ofte traff jeg eksempler som leksikalsk eller morfologisk sett var problematiske, men som syntaktisk sett var helt greie – som for eksempel den adjektiviske leddsetningen i eksempel 1, der studenten ikke hadde visst at subjunksjonen skal være *som*. I slike tilfeller der det bare var ukorrekt valgt subjunksjon har jeg markert setningen som syntaktisk korrekt, siden dette handler mer om manglende leksikalske ferdigheter. Hvis man derimot hadde utelatt subjunksjonen i en leddsetning der det ikke går å utelate, har jeg betraktet setningen som syntaktisk ukorrekt (som utelatt *at* i eks 2). På samme måte har jeg markert setningen ukorrekt hvis noen andre obligatoriske setningsledd mangler (som f.eks. manglende subjekt i den sideordnede setningen i eks 3).

Eksempel 1: *Små Thomas har et hund, hva heter Muki.* (student A10)

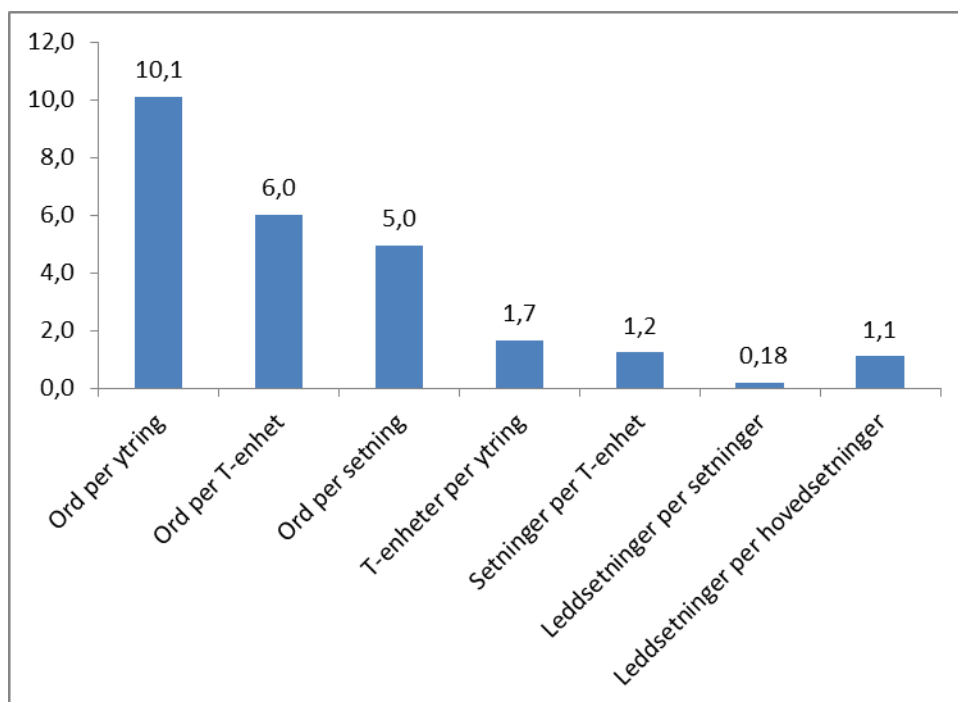
Eksempel 2: - *Hun håper frosk er i ugle hjemme, men ugle er ikke glad også.* (student A4)

Eksempel 3: *Hunden var også på vinduet men da flyet ut.* (student A3)

Som jeg har nevnt før, er den morfologiske, leksikalske og syntaktiske utviklingen gjennomvevd av hverandre. Dette må vi ha som baktanke når vi ser på den syntaktiske utviklingen, særlig kommer det til syne ved gruppe A. På grunn av det begrensede omfanget av en masteroppgave skal jeg konsentrere meg bare om den syntaktiske siden.

Vi skal begynne analysen av syntaktisk kompleksitet ved å se på det generelle nivået (figur 1). Som det framgår av figuret er T-enheter (6 ord) og setninger (5 ord) i gjennomsnitt nesten like lange i gruppe A, samtidig som de er nesten halvparten så lange som ytringer (10,1 ord). Dette viser at studentene bruker ganske mye sideordning mens underordning brukes mindre ofte. I gjennomsnitt består en ytring av 1,7 T-enheter, dvs. de fleste ytringer inneholder sideordning. En T-enhet inneholder som gjennomsnitt 1,2 setninger, som viser at de fleste T-enheter ikke inneholder noen underordning.

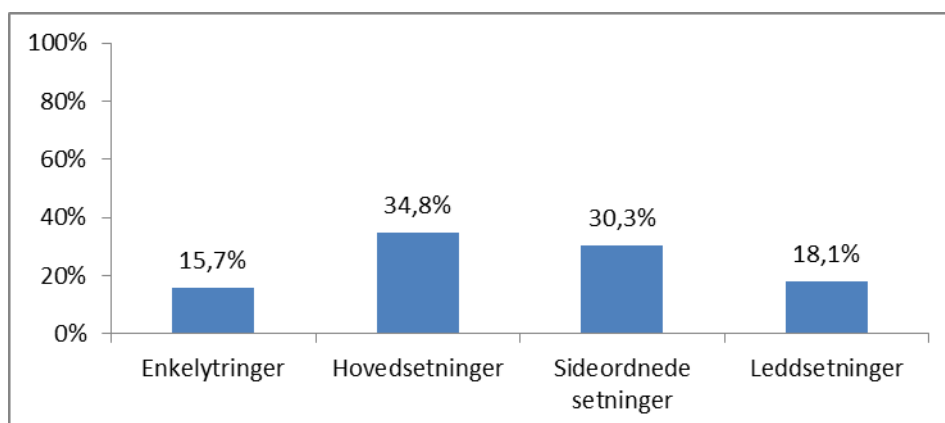
Figur 1: Gruppe A - generiske indikatorer



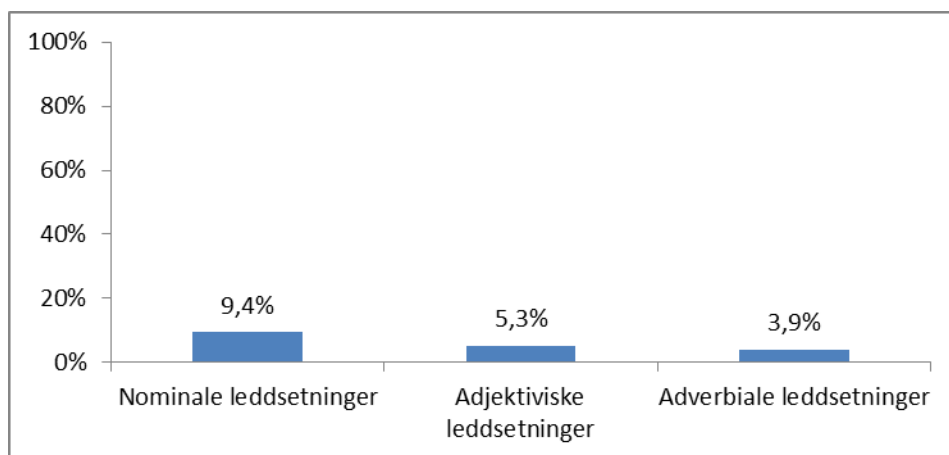
4.1.2 Hvilke strukturer brukes?

I figur 2 ser vi hvordan andelen av de forskjellige setningstyper fordeles prosentvis. Bare 16% av alle setninger er enkelytringer (i gjennomsnitt 6 ytringer per tekst), resten av ytringene kan betraktes som mer eller mindre komplekse strukturer: mer enn en tredjedel av setninger er hovedsetninger (ca 14 per tekst), dvs. de står sammen med enten en eller flere sideordnede eller underordnede setninger. Litt mindre enn en tredjedel av setninger er sideordnede setninger (ca 13 per tekst), samtidig som nesten en femtedel av setninger er leddsetninger (ca 7 per tekst).

Figur 2: Gruppe A - andelen av setningstyper (per alle setninger)



Figur 3: Gruppe A - andeler av leddsetningstyper (per alle setninger)



Tabell 7: Gruppe A - Andelen av underkategorier av leddsetninger (per alle leddsetninger)

Nominale leddsetninger...	56,2 %
- som objekt	56,2%
- som subjekt	0,0%
Adjektiviske leddsetninger...	26,3%
- ...med <i>som</i> som objekt	0,0%
- ...med <i>som</i> som subjekt	19,2%
- ..som tilhører et tidsuttrykk	5,5%
- ...som tilhører et stedsuttrykk	1,6%
Adverbiale leddsetninger	22,3%

Når det gjelder leddsetningstyper (se figur 3 og tabell 7), så er den vanligste typen av leddsetninger nominale leddsetninger, som utgjør omtrent halvparten av alle leddsetninger (ca 4 per tekst, 56,2%). Alle nominale leddsetninger fungerer som objekt i oversetningen (som i eksempel 4), mens det finnes ingen eksempel på nominal leddsetning som subjekt. De fleste nominale leddsetninger har oversetningen i form av «*X ser / hører / sier / tror / tenker at...*» .

Eksempel 4: *Da hørte Nils at froskene kvakker.* (student A7)

Størrelsesorden av adjektiviske og adverbiale leddsetninger er omtrent den samme: respektivt 26% (ca 2 per tekst) og 22% (ca 1-2 per tekst) av alle leddsetninger. I de fleste adjektiviske leddsetninger fungerer subjunksjonen *som* som subjekt i leddsetningen (19% av alle leddsetninger; se eks 5), derimot finnes det ingen adjektivisk leddsetning der *som* fungerer som objekt leddsetningen. Det forekommer både adjektiviske leddsetninger som tilhører tidsuttrykk (5,5%; eks 6) og de som tilhører stedsuttrykk (1,6%; eks 7).

Eksempel 5: *Hunden fant en bikube som falt ned av treet og hunden måtte løpe
vekk.* (student A5)

Eksempel 6: *En da natt, når gutten sov, gikk frosken ut av glasset.*
(student A5)

Eksempel 7: *Da de går til landskap og ser en bikube hver er mange bier og
vepser.* (student A8)

Adverbiale leddsetninger brukes for det meste for å uttrykke tid (eksempel 8) og sted (eks 9), i noen tilfeller også årsak (eks 10). I gruppe A kan man legge merke til ganske stor forvirring når det gjelder valg av subjunksjon (f.eks *hvor* i stedet for *der* eller *dit* for eksempel). Dette har jeg betraktet som manglende leksikalsk kunnskap (ikke syntaktisk) og markert setningene slik som jeg har forstått dem ved å støtte meg til konteksten. Adverbiale leddsetninger forekommer i oversetningen i både setningsinitial (eks 8) og setningsfinal (eks 9 og 10) posisjon. Den setningsinitiale plasseringen fører ofte til mislykket inversjon i oversetningen (se diskusjonen om syntaktisk korrekthet nedenfor).

Eksempel 8: *Når gutten og hunden sto opp, de funnet at der er ikke frosken.* (student A6)

Eksempel 9: *Først ser en hund mange bier og går hvor bier går.* (student A4)

Eksempel 10: *Der leket hunden med bikuben men det gjør vondt fordi der vært mange bier inne.* (student A11)

Selv om bare 18% av alle setninger er leddsetninger, synes jeg at det er merkelig at studenter på begynnernivå allerede prøver seg på de mer krevende syntaktiske strukturer. Leddssetninger finnes i alle 11 tekster og sammenlagt finnes det eksempler på alle typer leddsetninger unntatt adjektivisk leddsetning der *som* fungerer som objekt og nominal leddsetning som fungerer som subjekt (men disse typene var også veldig lite frekvente i de andre grupper, jf senere i kapittel 4.2.2 og 4.3.2.). Dessuten hadde man ikke behandlet disse strukturer i den formelle undervisningen ennå.

Forekomsten av inversjon i setninger er også nevneverdig frekvent – en femtedel av T-enheter (ca 7 per tekst; 20,9%) begynner med et annet setningsledd enn subjekt. Det setningsinitiale setningsleddet kan være både et enkelt ord (eks 11 og 13), en frase (eks 12) eller en leddsetning (eks 14) – som regel er det et adverbial.

Eksempel 11: *Plutselig er de i vann og hjorten kan ikke hjelpe dem.* (student A2)

Eksempel 12: *I bak av steinet var en hjort.* (student A5)

Eksempel 13: *Da vil han klimber på fjellet og hvis han går opp han ser en hjort som er veldig stor.* (student A8)

Eksempel 14: *Hvis gutten søver, frosken går ut på fiskeflaske/ glasse/ og går ut for å se landet.* (student A8)

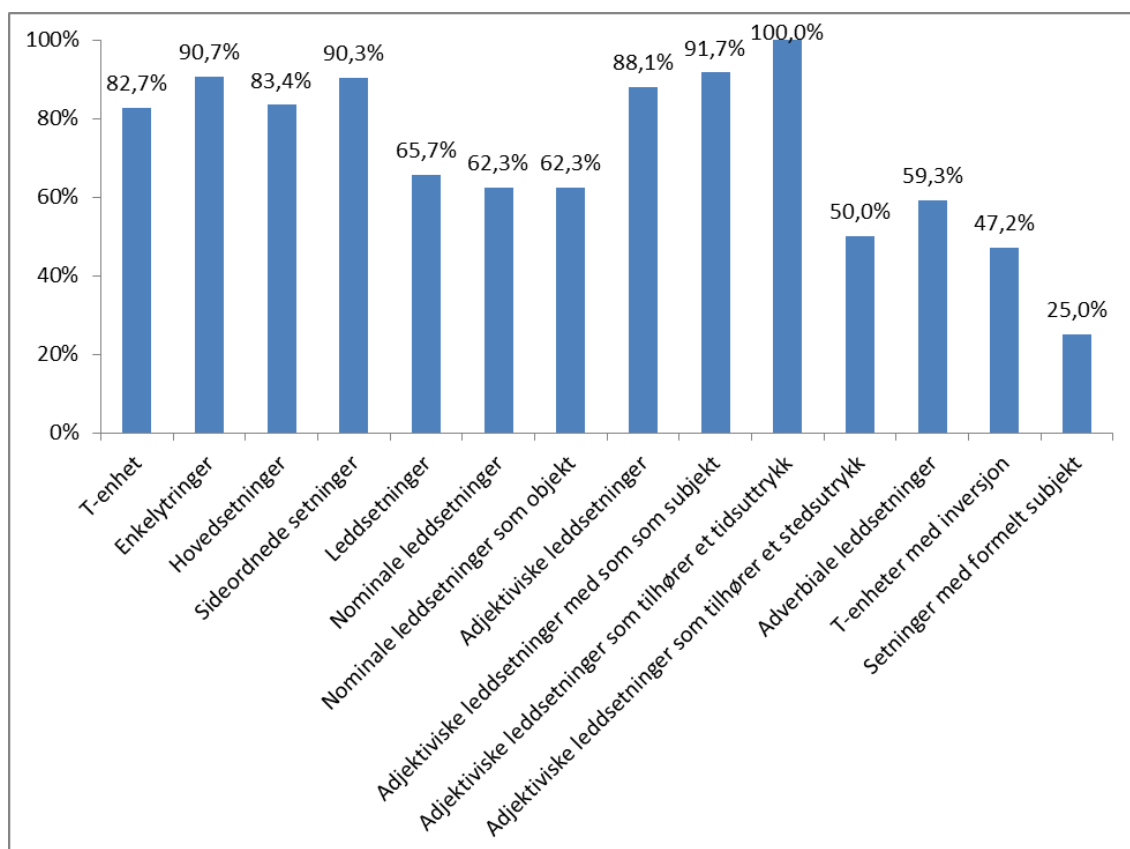
Når vi ser på de to foregående eksemplene fra student A8 (eks 13 og 14), ser vi at hun har brukt både sideordning og underordning, og har prøvd seg (ikke helt framgangsrikt men allikevel) på både adverbiale (*hvis han går opp / Hvis gutten søver*) og adjektiviske (*som er veldig stor*) leddsetninger og topikalisering av adverbialet (og derved inversjon). Selv om setninger ikke er helt korrekte og studenten sliter med mange språklige (grammatiske og leksikale) områder, kan vi legge merke til at setningsstrukturen hun har prøvd å lage er ganske innviklet. Dette kan sannsynligvis skyldes, at vi har å gjøre med voksne innlærere, som alle har høy kompetanse i sitt morsmål og har lært andre fremmedspråk før. Det vil bare være naturlig at de ikke vil uttrykke seg med kun enkelytringer og at behovet for å uttrykke mer komplekse tankeganger (selv om dette bare er en barnefortelling) fører til mer komplekse syntaktiske strukturer. Vi kan også se at andre fremmedspråk blir brukt som et ekstraressurs i tilfeller hvor norskkunnskapene ikke strekker til – det er tydeligst på leksikalsk nivå (f. eks *klimber* i eks 13), men brukes også på syntaktisk nivå. Den syntaktiske påvirkningen er vanskeligere å fastslå så klart, siden strukturene som er like i de respektive språk resulterer i korrekt syntaks på norsk. Således kan vi med større sikkerhet snakke om påvirkning fra andre språk bare der det finnes forskjeller mellom språkene som resulterer i ukorrekte syntaktiske strukturer på norsk (se diskusjonen og eksemplene 25-27 for noen av de syntaktiske trekkene nedenfor).

4.1.3 Syntaktisk korrekthet

Vi kan se at selv om studentene har forsøkt seg på forskjellige mer komplekse strukturer, så lykkes de ikke alltid å formulere dem korrekt. Figur 4 viser oss andeler av de syntaktisk korrekte strukturer i de respektive analysekategorier. På et generelt plan kan vi si at syntaktisk korrekthet i gruppe A ligger på et ganske høyt nivå. Litt mer enn 80% av alle T-enheter er syntaktisk sett helt riktige, for enkelytringer (n=65) og sideordnede setninger (n=147) er korrekthetsprosent enda høyere – litt mer enn 90% for begge. Leddsetninger (n=79) har vært en av de mest utfordrende kategorier for studenter, da bare 66% av dem er feilfrie. Samtidig finnes det noen typer av leddsetninger som ikke synes å skape problemer

– for eksempel adjektiviske leddsetninger som tilhører tidsuttrykk. Her må vi allikevel se litt mer bak statistikken: denne typen setninger forekommer veldig sjeldent, bare 4 ganger (studenter A5 og A10). Dermed sier disse 100% ikke at alle i gruppen skulle være i stand til å danne disse strukturer, men samtidig viser det at det på begynnernivå allikevel dannes slike strukturer og til og med feilfritt.

Figur 4: Gruppe A - andelen av syntaktisk korrekte strukturer



Nominale og adverbiale leddsetninger (hhv 62% (n=45) og 59% (n=17) korrekte) synes å være vanskeligere enn adjektiviske leddsetninger (88% (n=20) korrekte). I de fleste ukorrekte nominale leddsetninger har studentene problemer med plasseringen av setningsadverbialet (eks 15, 16, 17). Et annet hyppig kjennetegn er forsøk på å bruke inversjon i leddsetningen ved å plassere adverbialet først i setningen (eks 16, 17) .

Eksempel 15: *I morgen så gutt at frosken var ikke der.* (student A7)

Eksempel 16: *De ser at ganske frosken har hoppet ut av vinduen og hunden også hopper ut, det er godt, fordi nå han er fri fra glasset.* (student A1)

Eksempel 17: *Muldvarpen #er# ser at han vet ingenting og sier at kanskje uglen kan fortelle #whor# /hvor/ er frosken.* (student A2)

Setningsleddenes rekkefølge i leddsetningen er ikke helt klart heller. Det finnes eksempler på finitt verb foran subjektet i leddsetningen (eks 18).

Eksempel 18: *En gutt og en hund går på tur og ser hvor er en frosk.* (student A4)

Samtidig finnes det studenter som klarer eksakt samme type setning helt fint:

Eksempel 19: *Han og hunden hans gikk å se hvor frosken er.* (student A11)

En mulig forklaring her er at den setningstypen kan i seg selv også ses til å inneholde en slags inversjon, siden subjunksjonen i form av spørreordet *hvor* kan tolkes som et adverbial (*Hvor er frosken? Frosken er her.*), slik at studenten A4 (eks 18) i så fall bare har forsøkt å lage en invertert setning med hovedsetnings leddrekkefølge.

Ellers ser det ut at nybegynnere godt klarer å formulere leddsetninger som ikke inneholder setningsadverbial eller hvor de ikke har prøvd å lage inversjon (eks 21). Dette kan være med til å forklare hvorfor adjektiviske leddsetninger ikke har skapt mye vansker. Disse setninger var som regel korte og relativt enkle: det fantes kun én adjektivisk leddsetning i gruppe A som inneholdt nektelse (og den var også syntaktisk ukorrekt, se eks 20) eller noe annet setningsadverbial, ikke heller fantes det forsøk på å lage inversjon i adjektiviske leddsetninger.

Eksempel 20: *Der vær også en hjørt som vær ikke så god.* (student A11)

Eksempel 21: *I bak av et gamle treet, som hadde falt ned ved siden av vatnet, fant gutten og hunden frosken.* (student A5)

Inversjon og formelt subjekt synes å være de mest kompliserte strukturer som studentene i gruppe A ikke helt har tilegnet seg ennå. Inversjon forekom ganske ofte (i 21% av alle T-enheter), mens mindre enn halvparten (47%) av T-enheter som inneholder inversjon er korrekte (se figur 4). Ved å se nærmere på tekstene, finner vi stort sett to typer setninger der inversjon skaper problemer: setningene der adverbialet har blitt flyttet til Forfeltet (eks 22, 23) og setninger der Forfeltet inneholder en leddsetning som fungerer som adverbial (eks 24).

Eksempel 22: *Først han se at Kvakk er ikke i glasset.* (student A9)

Eksempel 23: *I morgen gutten og hunden stå opp og så at frosken #var# var ikke i glasset.* (student A3)

Eksempel 24: *Kvakk liker ikke å bo i glasset og når Jan og Rex sovet, Kvakk #gå# gikk ut.* (student A9)

Som nevnt før, finnes det også forsøk på å bruke inversjon i en leddsetning (eks 25). I slike tilfeller har jeg allikevel markert det som bruk av inversjon (selv om det ikke er korrekt og at det i prinsipp aldri kan forekomme inversjon i leddsetningen). Det ser ut som om dette kan være en påvirkning fra estisk, siden vi ikke har en slik restriksjon for estiske leddsetninger og det høres helt naturlig ut å begynne en leddsetning med et adverbial på estisk (jf med eks 25 – «...sest nüüd sai ta klaasist lahti.»)

Eksempel 25: *De ser at ganske frosken har hoppet ut av vinduen og hunden også hopper ut, det er godt, fordi nå han er fri fra glasset.* (student A1)

De ovenstående eksemplene (eks 22-25) illustrerer hvordan syntaktisk kompleksitet kan hope seg opp slik at det er flere syntaktiske oppfordringer i samme ytring som må takles samtidig (som f. eks inversjon, sideordning, underordning).

Jeg tok en liten kikk i læreboken som ble brukt i undervisningen med nybegynnergruppen (Ellingsen og Mac Donald 2004) og så at V2 regelen introduseres allerede i første kapittel, til og med i forbindelse med inversjon. Dermed har vi å gjøre med et syntaktisk trekk som studentene skulle ha lært seg ganske tidlig. Derimot tar man opp inversjon i forbindelse med komplekse strukturer (for eksempel en foranstilt leddsetning) først ved slutten av læreboken. Ved innsamlingstidspunktet av min materiale hadde studentene ikke kommet så langt i læreboken, samtidig prøvde de allerede på inversjon ved mer komplekse setninger. Dette viser at undervisningen skulle inkludere mer syntaktisk kompleksitet allerede tidligere.

Det andre ganske problematiske området var bruk av formelt subjekt. Setninger med formelt subjekt var korrekt dannet i bare 25% (n=23) av tilfeller. Her må jeg allikevel kommentere forskjellen mellom den kategorien og andre. Jeg la merke til en hel rekke presenteringssetninger - dvs. setninger som skal begynne med et formelt subjekt og der det egentlige subjektet inneholder ny informasjon og står i ubestemt form – for eksempel *Det sitter ei jente på trappa* (Golden *et al.* 2011: 157). I kontekster der det presenteres ny informasjon (i vår tilfelle for eksempel da det kommer en ny karakter inn i historien), er det om ikke 100% obligatorisk, da det mest naturlige på norsk å bruke en presenteringssetning med formelt subjekt. I gruppe A og også ved et par tilfeller i gruppe B

var slike setninger ofte formet uten formelt subjekt. I slike tilfeller har jeg valgt å markere setningen som ukorrekt. Ofte kunne man i den slags setningen se innvirkning fra enten andre fremmedspråk eller morsmålet. Det var flere tilfeller der studenten i stedet for å bruke presenterende konstruksjon «*Det er / var*» begynte setningen med adverbialet «*Der*» som i eksemplene 26 og 27.

Eksempel 26: *Der vær en liten gutt som heter Fredrik.* (student A11)

Eksempel 27: *Imorgen gutten og hunnen vekker, de står #o#upp og ser, at /der er ingenting i/ glasset.* (student A1)

Her kan det diskuteres om man har brukt inversjon og flyttet adverbialet foran i setningen eller om det kanskje er påvirkning av den engelske presenteringsformuleringen «*there is nothing in the jar*». Hvordan det enn er, er det klart at det skulle være et formelt subjekt i den norske setningen. Dermed speiler den lave korrekthetsprosenten for formelt subjekt (25%) både setninger der det var et formelt subjekt i setningen men som av en eller annen grunn var syntaktisk ukorrekte, samt setninger der det skulle ha vært et formelt subjekt på norsk, men studenten hadde ikke brukt det.

Det at bruk av formelt subjekt ikke kommer helt naturlig til studentene skyldes sannsynligvis også faktumet at deres morsmål estisk ikke har et slikt kategori. Ved materialets innsamlingstidspunkt hadde studentene ikke fått noen direkte undervisning i bruk av formelt subjekt. Siden dette ikke heller er noe som kan overføres fra morsmålet, ble korrekthetsprosenten lav i den kategorien.

Et interessant funn ved materialet var at studentene bruker en slags «gjenvinning» av strukturer – bruk av samme struktur i litt forskjellige kontekster. Man kan klart se at visse studenter har brukt visse strukturer mer enn de andre. Student A5 for eksempel bruker flittig adjektiviske leddsetninger med *som* som subjekt (til og med i samme ordlyd – eks 28 og 30), mens student A10 bruker adjektiviske leddsetninger som tilhører tidsuttrykk og A8 adverbiale leddsetninger.

Student A5:

Eksempel 28: *„Frosken, er du her?“, men der bodde #ett# en #mol# muldvarp, som skremte gutten.*

Eksempel 29: *Hunden fant en bikube som falt ned av treet og hunden måtte løpe vekk/.*

Eksempel 30: *Gutten prøvde å finne frosken i treet, men der bodde en ugle, som skremte gutten og gutten falt ned av treet.*

Student A10:

Eksempel 31: *Om natt, når Thomas og Muki sover, kommer frosken ut fra glassen.*

Eksempel 32: *I morgen, når de står opp, de ser ikke frosken.*

Student A8:

Eksempel 33: *Hvis gutten sover, frosken går ut på fiske#flaske#/glasse/ og går ut /for/ å se landet.*

Eksempel 34: *Hvis de #går# /står/ opp, de ser at frosken har gikk ut.*

Eksempel 35: *Da vil han #?#klimber på fjellet og hvis han går #?# opp han ser en hjort #?# som er veldig stor.*

Samtidig kan man se fra eksemplene 28-35 at mellomspråkssyntaks kan være ganske stabil og regelmessig hos noen informanter og mer dynamisk og ustabil hos de andre. Ved samme type setningsstruktur lykkes de å lage både korrekte og ukorrekte varianter. Student A8 viser seg å være ganske stabil i ytringene der hun bruker adverbiale leddsetninger: i alle forannevnte eksempler (33-35) klarer hun og lage syntaktisk korrekte leddsetninger, samtidig som hun i alle hovedsetninger plasserer subjektet foran det finite verbet (dvs. hun har ikke klart å lage inversjon korrekt). Student A10 synes derimot ikke å ha noen fast syntaktisk regel når det gjelder inversjon. Eksemplene 31 og 32 er helt identiske i sin struktur, allikevel inverterer hun en av hovedsetningene helt korrekt, mens hun i den andre bruker normal ordstilling i stedet for inversjon.

For å sammenfatte resultatene for gruppe A, kan vi si at voksne studenter på begynnernivå allerede bruker ganske kompliserte syntaktiske strukturer, selv om deres ordforråd eller grammatiske ferdigheter ikke er på særlig høyt nivå. De lykkes ikke alltid å skape syntaktisk korrekte setninger og sliter med mange syntaktiske kategorier, men samtidig finnes det også flere kategorier som de har mestret på ganske høyt nivå.

4.2 Viderekommende gruppen (Gruppe B)

4.2.1 Generelt om tekstene

Materialet for gruppe B omfatter 10 tekster. Følgende tabell gir oversikt over de viktigste karakteristika av tekstene og gruppen:

Tabell 8: *Tekstene i Gruppe B*

		Maksimum	Minimum
Antall tekster	10		
Antall ord sammenlagt	4676		
Ord per en tekst	468	800	259
Antall T-enheter sammenlagt	594		
T-enheter per en tekst	59		
Antall setninger sammenlagt	823		
Setninger per tekst	83		
Antall ytringer sammenlagt	602		
Ytringer per tekst	39		

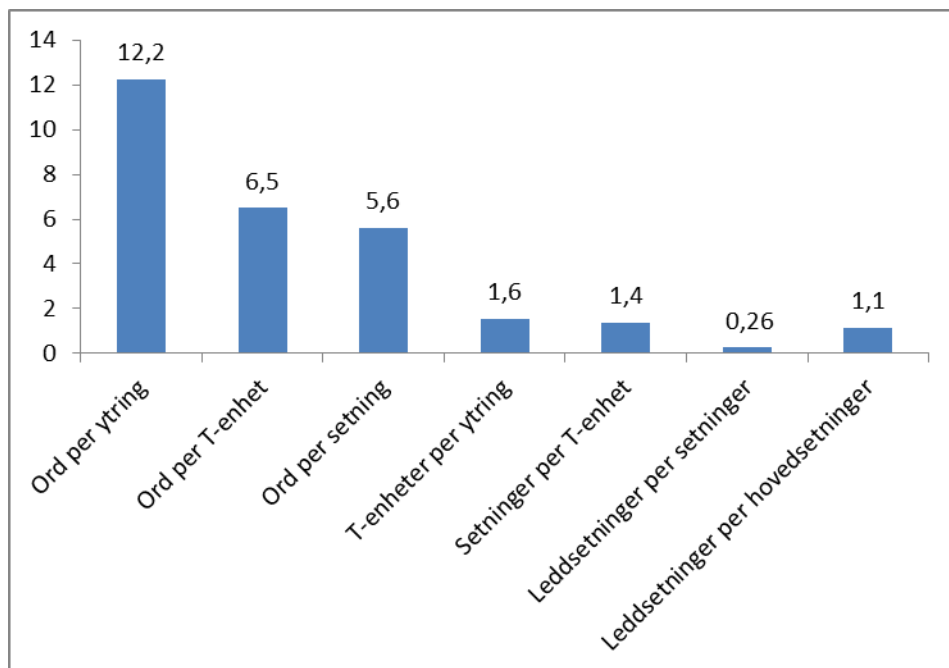
Tekstene i gruppe B er lengre og grundigere, studentene beskriver bildefortellingen i detalj og de fleste tar nøyaktig med hver eneste bilde. Som vi kan se i tabell 8 er variasjonen i gruppen ganske stor, det finnes noen kortere tekster og et par veldig lange tekster. Medianverdiene ligger derfor litt lavere enn gjennomsnittsverdier i alle kategorier: antall ord per tekst 424, T-enheter per tekst 54, setninger per tekst 76 og ytringer per tekst 35. Jo mindre syntaktisk enhet vi betrakter, jo mindre blir forskjellen mellom gjennomsnittsverdiene og medianer. Med hensyn til bedre sammenlignbarhet med andre grupper (der jeg har presentert gjennomsnittsverdier), har jeg brukt ved analysen av gruppe B bare gjennomsnittsverdier – ved analysering av de forskjellige syntaktiske strukturer sjekket jeg at forskjellene mellom median og gjennomsnitt ikke var særlig store.

4.2.2 Hvilke strukturer brukes?

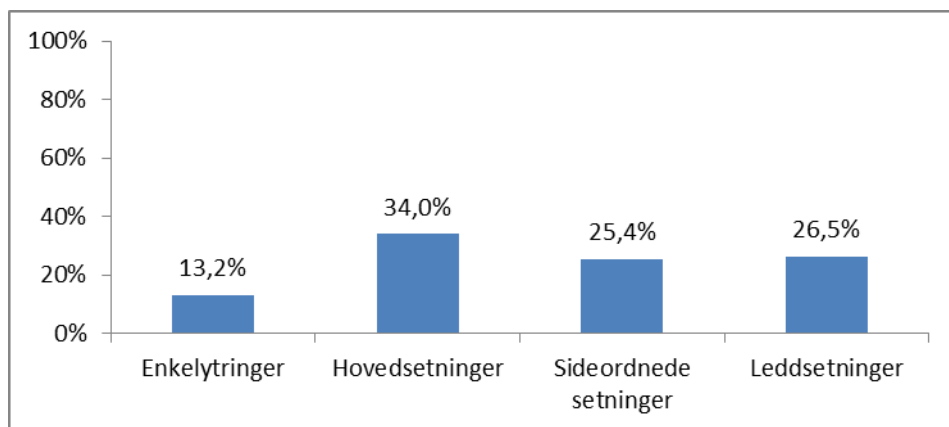
Når vi ser på det generelle nivået av syntaktiske indikatorer (figur 5), ser vi at i gruppe B er ytringer (12,2 ord) i gjennomsnitt omtrent dobbelt så lange som T-enheter (6,5 ord) og setninger (5,6 ord). Dette viser igjen mot lavere frekvens av enkelytringer (ca 11 per tekst; 13% av alle setninger) og høyere frekvens av sideordning (21 per tekst; 25%) og underordning (22 per tekst; 27%). En tredjedel av alle setninger er hovedsetninger (28 per

tekst; 34%) som har enten en (eller flere) sideordnet eller underordnet konstruksjon tilknyttet. Bruk av sideordning og underordning er ganske likt i gruppe B: sideordnede setninger utgjør 25% av alle setninger og leddsetninger 27%. (se figur 6)

Figur 5: *Gruppe B - generiske indikatorer*

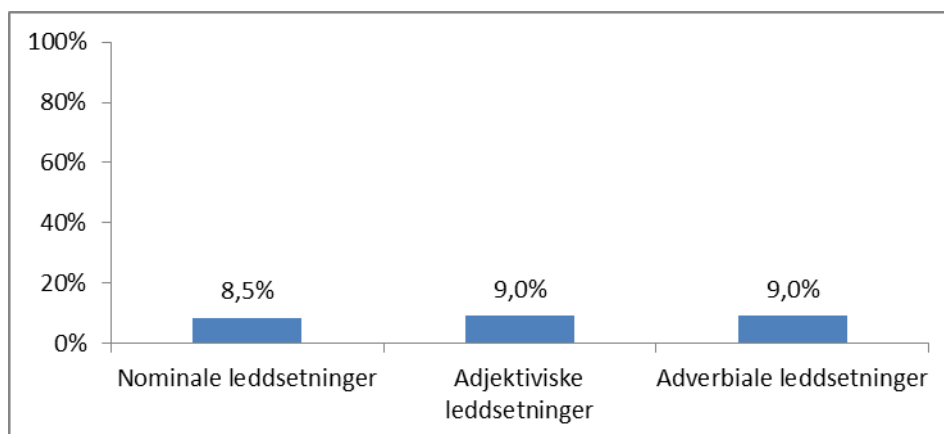


Figur 6: *Gruppe B - andelen av setningstyper (per alle setninger)*



Fordelingen av de forskjellige typer leddsetninger (figur 7) er også ganske lik. Studentene bruker omtrent like mange nominale (7,1 per tekst; 8,5% av alle setninger), adjektiviske (6,7 per tekst; 9%) og adverbiale (7,9 per tekst; 9%) leddsetninger. Inversjon er igjen relativt hyppig brukt: i nesten en fjerdedel av alle T-enheter (ca 14 per tekst; 24%). Formelt subjekt brukes i 7% av alle setninger (ca 5 per tekst).

Figur 7: Gruppe B - andeler av leddsetningstyper (per alle setninger)



Tabell 9: Gruppe B - Andelen av underkategorier av leddsetninger (per alle leddsetninger)

Nominale leddsetninger...	33,6%
- som objekt	33,3%
- som subjekt	0,4%
Adjektiviske leddsetninger...	35,5%
- ...med <i>som</i> som objekt	5,6%
- ...med <i>som</i> som subjekt	21,5%
- ..som tilhører et tidsuttrykk	2,3%
- ...som tilhører et stedsuttrykk	4,3%
Adverbiale leddsetninger	34,1%

Når vi ser på de forskjellige underkategorier av leddsetninger (tabell 9), så ser vi at alle typer leddsetninger er representert i gruppe B. Omtrent en tredjedel av alle leddsetninger er nominale leddsetninger som fungerer som objekt i oversetningen (33,3%). Det fantes også ett eksempel på nominal leddsetning som subjekt og dette var til og med korrekt formulert (eks 36).

Eksempel 36: *Og det aller viktigste var det at de hadde fanget en frosk.* (student B6)

Dette er i og for seg en underkategori av nominal leddsetning som subjekt – slik at leddsetningen fungerer her som egentlig subjekt, ettersom vi har et formelt subjekt i setningen.

De fleste adjektiviske leddsetninger har *som* som subjekt (21,5%) i leddsetningen, men det fantes også adjektiviske leddsetninger der *som* fungerer som objekt (5,6%). I flere steder hadde studentene utelatt subjunksjonen *som* fra leddsetningen, slik som man godt kan gjøre på norsk hvis *som* ikke er subjekt i setningen (eks 37, 38).

Eksempel 37: *Gutten fortsatt med å lete etter frosken og han sa i et hylle han fant i et alt tre.* (student B2)

Eksempel 38: *Hjemmelengsel var noe han kjente selv og Roald lovte Mala, at hun kan besøke familien hennes, hvis hun vil.* (student B4)

I eksemplet 38 kan man allerede se en ganske innviklet syntaktisk struktur med kombinasjon av både sideordning og underordning, samt forskjellige typer leddsetninger som til og med fungerer som oversetninger for en annen leddsetning (andregradsunderordning).

Det forekom også noen få eksempler på andre typer adjektiviske leddsetninger – de som tilhører tids- (2,3%) og stedsuttrykk (4,3%):

Eksempel 39: *Roald hadde funnet et hull i jorden og tenkte, at det kunne være akkurat det stedet hvor dårlige menn hadde lagt Mala.* (student B4)

Eksempel 40: *Det er da de ser hvor frosken hadde forsvunnet.* (student B10)

Eksempel 41: *Både gikk mot plassen kvaking kom fra og de sa etter et tre som hadde falt ned.* (student B2)

Som vi ser fra de siste to eksemplene, så fantes det også her setninger der subjunksjonen hadde blitt utelatt ettersom den ikke fungerte som subjekt i setningen (syntaktisk korrekt på norsk).

Variasjon av adverbiale leddsetninger (34,1%) var stor. Det forekom ikke bare adverbiale leddsetninger som uttrykker tid eller sted, men også andre mer komplekse betydninger som motsetning (eks 42), årsak (eks 43) eller sammenlikning (eks 44).

Eksempel 42: *Selv om frosken likte det med an??e froske<n/r>, ville han gå tilbake hjemme.* (student B2)

Eksempel 43: *Jeg ropte etter den først uten å gå inn i skogen, ettersom jeg har aldri vært der.* (student B7)

Eksempel 44: *Frosken satt der i glasset og så#g# ut /slik/ som ingenting hadde skjedd.* (student B7)

Selv om det ikke har noen betydning for den syntaktiske kompleksiteten hvilken type adverbial leddsetning man bruker, kan vi se at da repertoaret av subjunksjonene som tilhører adverbiale leddsetninger er bred (slik som det er for gruppe B), så bruker studentene den typen setninger ganske hyppig. Altså det er igjen vanskelig å skille den syntaktiske utviklingen fra den leksikalske.

4.2.3 Syntaktisk korrekthet

Hva angår den syntaktiske korrektheten, er resultatene fra gruppe B veldig høye (se figur 8). Alle kategoriene, unntatt én (formelt subjekt), inneholder mer enn 90% syntaktisk korrekte setninger. I gjennomsnitt er 95% av alle T-enheter syntaktisk korrekte.

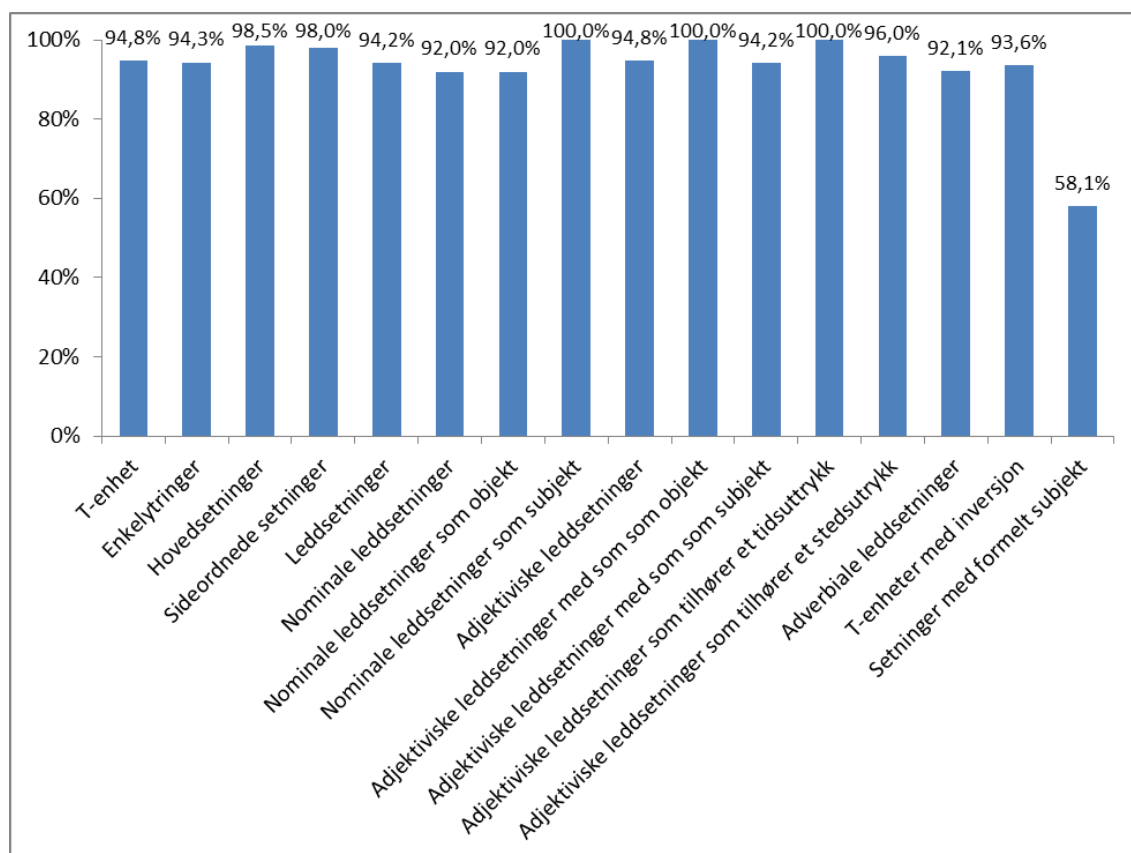
Hovedsetninger (98,5%) og sideordnede setninger (98%) er hovedsakelig korrekt formulerte, andelen av korrekte setninger er litt lavere for enkelytringer (94,3%) og leddsetninger (94,2%). I enkelytringene er det mest påfallende problemet at man ikke bruker formelt subjekt der det skal brukes (dette skal jeg kommentere nærmere på litt senere):

Eksempel 45: *Der bodde en ugle.* (student B3)

Eksempel 46: *Etter en liten tid kom sju f#o#rosker igjen.* (student B3)

Eksempel 47: *Da var hans gang å løpe for sitt liv.* (student B4)

Figur 8: Gruppe B - andelen av syntaktisk korrekte strukturer



På lik linje med gruppe A kan vi se at enkelytringer (eks 45-47) er ikke alltid enkle i sin setningsstruktur og kan by på flere utfordringer i seg selv. De kan inneholde inversjon og/eller formelt subjekt eller være utvidet på frasenivå (ved hjelp av attributter, infinitte underordninger osv). Dette vil skape et interessant diskusjon om vi kan i det hele tatt betrakte en høy andel av enkelytringer som et tegn på lavere kompleksitetsnivå. På den ene siden henviser en høy andel enkelytringer til lavere bruk av sideordning og underordning, på den andre siden kan den setningsinterne kompleksiteten allikevel være på høyt nivå.

Når vi ser på resultatene av undertypene av leddsetninger, så finnes det til og med 3 kategorier der alle setninger var korrekt formulerte: nominale leddsetninger som subjekt, adjektiviske leddsetninger som tilhører et tidsuttrykk og adjektiviske leddsetninger med *som* som objekt. De første to kategoriene er veldig lavfrekvente, det fantes bare 1 eksempel på nominal leddsetning som subjekt (eksempel 36 ovenfor) og 3 eksempler på adjektivisk leddsetning som spesifiserer tiden (eks 48-50). Selv om slike setninger ikke forekommer

ofte, ser man her ganske innviklede syntaktiske konstruksjoner, med innskutte leddsetninger som er plassert mitt i oversetningen, samt utelatt subjunksjon:

Eksempel 48: *Om morgenen, da #gutten# /han/ våknet, oppdaget gutten at #frosken# Bjarne hadde gått bort.* (student B5)

Eksempel 49: *Neste morgen#en# da gutten og hunden stod opp, ble de sjokkert.* (student B6)

Eksempel 50: *Det er da de ser hvor frosken hadde forsvunnet.* (student B10)

Adjektiviske leddsetninger med *som* som objekt forekom derimot litt hyppigere (10 ganger). Jeg har betraktet adjektiviske leddsetninger der *som* fungerer som objekt mer komplekse enn de der *som* fungerer som subjekt. I den andre typen har vi vanligvis å gjøre med leddsetningsstruktur SVO, der *som* har tatt plassen for S. I den første typen derimot har vi i leddsetningen OSV struktur og subjunksjonen fungerer her som O – dermed kan dette fortolkes som foranstilt objekt, samtidig som på grunn av ordstillingen i leddsetningen vi ikke kan bruke inversjonsregelen. Således har vi å gjøre med en ganske innviklet konstruksjon. Til tross for det, var alle forekomster av adjektiviske leddsetninger der *som* fungerer som objekt korrekt formet i gruppe B (eks 51). Dette kan derfor ses som et relevant tegn på et høyt nivå av den syntaktiske utviklingen.

Eksempel 51: *I steden fo?r den frosken han hadde før, tok han en liten barnefrosk med seg.* (student B2)

I seg selv er ytringen i eksempel 51 ikke særlig lang, samtidig som kompleksiteten kommer til uttrykk gjennom et foranstilt adverbial som i sin tur inneholder en adjektivisk leddsetning der subjunksjonen er utelatt.

Kategorier som har litt lavere korrekthetsgrad var nominale leddsetninger som objekt (92%) og adverbiale leddsetninger (92,1%). I de nominale leddsetningene oppsto feilene igjen på grunn av at man ikke hadde brukt formelt subjekt i setningene der man skulle ha gjort det (eks 52, 53). Oftest hadde man også flyttet adverbialet til begynnelsen av leddsetningen som det ikke går an å gjøre på grunn av den faste leddstillingen i norske leddsetninger (eks 53, 54).

Eksempel 52: *Vi bestemte å lete etter frosken #?#/f/ra /skogen, fordi vi visste at mange dyr bor der.* (student B7)

Eksempel 53: *De forstår at der er også damefrosk, og de har flere barn.* (student B9)

Eksempel 54: *De klatrer over et bjelke, og ser at her står frosken de hadde lett etter.* (student B9)

Det fantes også eksempler hvor plasseringen av setningsadverbialet ble galt:

Eksempel 55: *Men midt i natten bestemte frosken at #ha#/de/n #skal# vil ikke #være# sitte i glasset med #ingenting# /ikke noen/ å gjøre, og derfor flyktet den.* (student B7)

Det mest problematiske området når det gjelder syntaksen i gruppe B ser allikevel ut til å være formelt subjekt. Bare litt mer enn halvparten (58%) av alle setninger var korrekte i den kategorien. Her må jeg igjen påpeke, at jeg har markert setningen feil også da det ikke fantes formelt subjekt i setningen men det skulle ha vært der. I setningene der det formelle subjektet fantes, oppstod problematiske momenter ved noen andre syntaktiske områder, for eksempel ved plasseringen av setningsadverbialet i leddsetningen:

Eksempel 56: *Jens var litt sint, men samtidig også lykkelig at det skjedde ikke noe vondt med hunden.* (student B1)

I presenteringssetningene der det formelle subjektet ikke fantes, forekom det (på samme måte som i gruppe A) ganske ofte setningsinitialt bruk av adverbialet *der* som peker igjen mot påvirkning av engelsk «*There is / are*».

Eksempel 57: *Der var bare en muldvarp som likte det ikke at noen ropte i hjemmet hans.* (student B2)

Eksempel 58: *De forstår at der er også damefrosk, og de har flere barn.* (student B9)

I noen tilfeller ville jeg tolke mangelen av det formelle subjektet som påvirkning av morsmålet estisk:

Eksempel 59: *Der bodde en ugle som var forstyrret av det.* (student B1)

Eksempel 60: *Da var hans gang å løpe for sitt liv.* (student B4)

Eksempel 61: *Når gutten står på et stort stein og roper etter frosken, kommer bak steinen en hjort som hjelper å bære gutten på seg, slik at de kan gå fortere.* (student B9)

På norsk skulle alle de understrekte setninger inneholde et formelt subjekt, noen av dem skulle til og med bli omformulert på grunn av feil leddstilling. En direkte omformulering av setningene til estisk, ville derimot lyde helt naturlig, for eksempel «*Seal elas üks öökull...*» eller «*Siis oli tema kord joosta...*».

Samtidig finnes det også studenter som bruker formelt subjekt framgangsrikt og til og med hyppig. Student B6 for eksempel har 11 slike setninger (dette utgjør 9% av alle setninger i hans tekst) og alle disse er korrekte. Setningene hans låter veldig flytende og naturlige på norsk:

Eksempel 62: *Nå var det allerede langt på natt, og gutten og hunden var i soverommet.*

Eksempel 63: *Det var en hel froskfamilie bak treet.*

Eksempel 64: *Men det var en muldvarp som kom ut av hullet.*

Alt i alt kan vi se at studenter i gruppe B bruker en bred repertoar av forskjellige syntaktiske strukturer og syntaktisk kompleksitet er ganske høy. De fleste komplekse strukturer klarer de helt fint, andelen korrekte setninger er også svært høy. Det eneste problematiske området som ikke alle har tilegnet seg helt og som ikke kommer naturlig for studentene er bruk av formelt subjekt.

4.3 Sammenlikning av alle grupper (inkludert gruppe C: morsmålstalere)

4.3.1 Generelt om tekstene

Først skal vi se på de generelle trekk ved tekstene i alle grupper side ved side. Følgende tabell skal hjelpe oss å sammenlikne tekstene:

Tabell 10: *Tekstene i alle grupper*

	Gruppe A	Gruppe B	Gruppe C
Antall tekster	11	10	5
Antall ord sammenlagt	2240	4676	3046
Ord per en tekst	204	468	609
Antall T-enheter sammenlagt	377	594	330
T-enheter per en tekst	34	59	66
Antall setninger sammenlagt	454	823	480
Setninger per tekst	41	83	96
Antall ytringer sammenlagt	373	602	205
Ytringer per tekst	21	39	41

Jeg hadde ikke som hensikt at gjennomføre en egen undersøkelse av gruppe C – norsk som morsmål – men bare bruke den som en referansegruppe for de andre to grupper. Jeg har analysert tekstene til gruppe C ved hjelp av de samme analyseredskap og målestokk, men jeg skal ikke presentere funnene fra gruppe C separat. Derimot skal jeg i den delen av analysen sammenlikne mine resultat fra gruppene A og B seg imellom, samt sammenlikne dem med gruppe C.

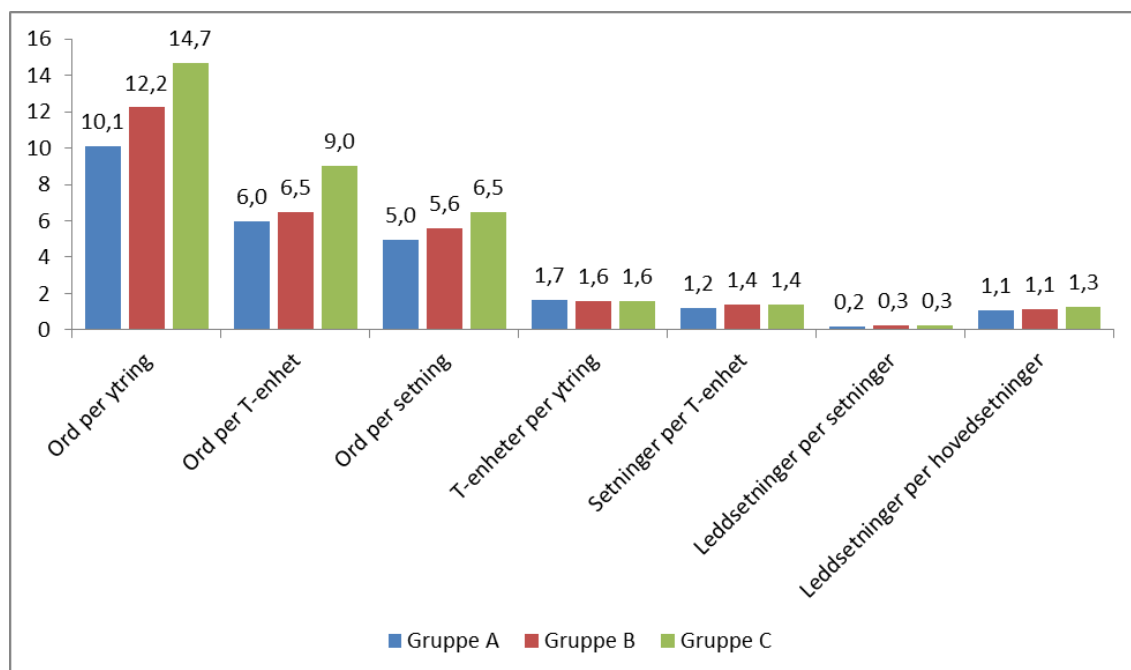
Som vi ser fra tabell 10, inneholder tekstene i gruppe B den største tekstmassen (4676 ord sammenlagt), samtidig er også de andre gruppene ganske godt representert for å trekke konklusjoner. Selv om det bare finnes 5 tekster fra morsmålstalende informanter, er tekstene deres mye lengre. Når vi ser på antallet T-enheter, setninger og ytringer per tekst, så er tekstene i gruppe B på lik linje med morsmålstekstene, mens nybegynnergruppen har skrevet mye kortere tekster. Siden ingen av gruppene fikk noen begrensninger på hvor lenge eller hvor mye de måtte skrive, avspeiler resultatene at gruppe A, logisk nok, fortsatt har begrensede språklige ferdigheter. Når vi ser nærmere på tekstene, legger vi klart merke til at gruppe A har vært mer lakonisk med sin beskrivelse av det som hender på bildene og sannsynligvis har de valgt å ta opp bare disse aspekter fra bildene som de følte at de har nok språklige ressurser til å forklare. Gruppe B var mer detaljert i sine beskrivelser og tok

med flere aspekter fra bildene, det fantes ganske fantasifulle historier der man selv hadde funnet på interessante nyanser ved fortellingen som ikke fantes på bildene.

Ved sammenlikning av de generelle indikatorer av syntaktisk kompleksitet (figur 9) viser det seg at både ytringslengde, lengden av T-enheten og setningslengden øker fra gruppe A til B og fra B til C. Samtidig er nivået av T-enheter per ytring og setninger per T-enhet nesten ens i alle grupper. Gruppe A skiller seg bare litt ut her, antallet T-enheter per ytring er litt høyere for dem (1,7), samtidig som setninger per T-enheter er litt lavere (1,2). Dette viser at sideordningsgraden for gruppe A er høyere og underordningsgraden lavere enn for de andre grupper. Forskjellene er allikevel ganske små.

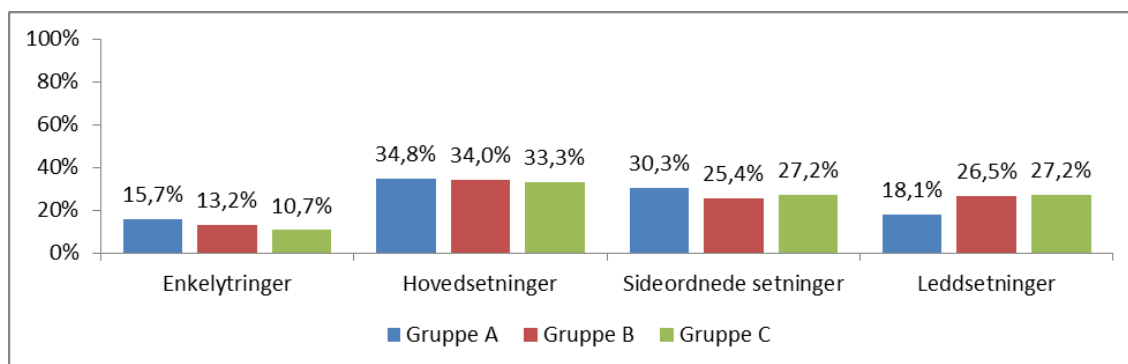
Dette peker mot noe som jeg også kommenterte i kapittel 2. – nemlig at vi ikke kan betrakte den syntaktiske kompleksiteten bare som bruk av leddsetninger og utviklingen av den syntaktiske kompleksiteten som et evig økende underordningsgrad. Språkbrukere som er på høyere nivå av språkbeherskelse bruker ikke nødvendigvis mer underordning eller sideordning eller større grad av underordning. Vi må også se på hvilken type syntaktisk komplekse strukturer som brukes og ikke minst hva som skjer setningsinternt. Hvis både lengden av ytringer, T-enheter og setninger er større uten at sideordning eller underordning øker, må det være noe innenfor setningen som gjør setningene lengre (se også analysen av de lengste syntaktiske enheter nedenfor).

Figur 9: Alle grupper - generelle indikatorer



Når det gjelder valg av de forskjellige syntaktiske strukturer og setningstyper (figur 10 og 11 og tabell 11), står gruppene ganske nære hverandre. Det finnes forskjeller, men størrelsesorden av resultatene er stort sett ens. Dette betyr at i sitt valg og variasjon av setningstyper likner studentenes mellomspråkssyntaks målspråket allerede ved begynnernivået, men enda mer på viderekommende nivået.

Figur 10: Alle grupper - andelen av setningstyper (per alle setninger)



Som det framgår av figur 10 kan vi betrakte en liten nedgang på andelen av enkelytringer i tekstene, dvs. begynnergruppen bruker litt flere enkelytringer (15,7%) i sine tekster enn den viderekommende gruppen (13,2%), som i sin tur har mer enkelytringer enn morsmålstekstene (10,7%). Det samme gjelder sideordning: gruppe A har litt høyere frekvens av sideordnede setninger (30,3%), mens gruppene B (25,4%) og C (27,2%) har litt mindre. Den største forskjellen ser vi allikevel i bruk av leddsetninger, der forskjellen mellom gruppe A (18,1%) og C (27,2%) er nesten 10%, samtidig som gruppe B (26,5%) er på lik linje med gruppe C (27,2%). Vi kan også se en liten økning i antallet leddsetninger per hovedsetninger (se figur 9) som er høyere for gruppe C (1,3) enn de andre grupper (1,1 for begge) – dette betyr at i gruppe C finnes det flere leddsetninger som er i sin tur underordnet en annen leddsetning (dvs. andre-, tredje- og fjerdegradsunderordnede setninger), som for eksempel:

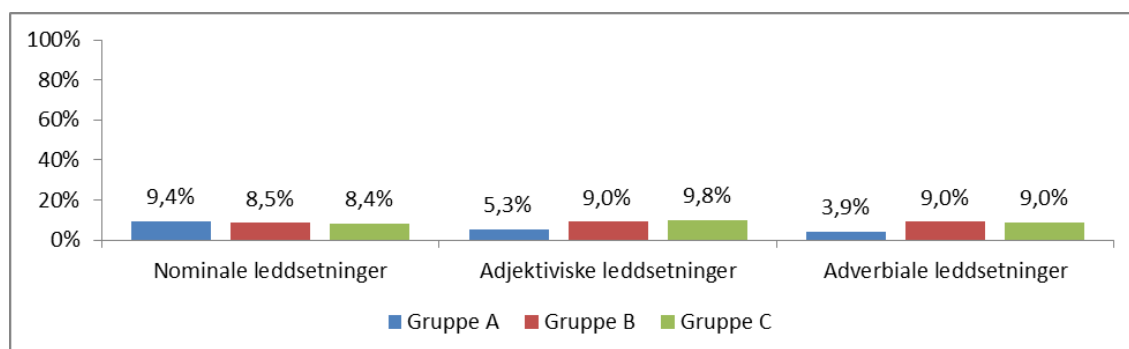
Eksempel 65: *Vepsebolet hadde ramla ned fra greina si, og ut kom en sverm av illsinte insekter som tydeligvis trodde (førstegradsunderordning) at hunden var skyld i (andregradsunderordning) at hjemmet deres ramla i bakken (tredjegradsunderordning).* (informant C1)

Slike underordninger finner vi også ved gruppe A og B (eks 66, 67, 70), men da har vi likevel å gjøre med maksimalt andregradsunderordning. Tredjegradsunderordning fantes ikke i materialet for hverken gruppe A eller B.

4.3.2 Hvilke strukturer brukes?

Figur 11 gir oss oversikt over bruk av de forskjellige syntaktiske underkategorier i forhold til alle setninger. Her kan vi se at de fleste kategoriene ser ganske ens ut, størrelsesorden for bruk av de forskjellige kategoriene er stort sett den samme for alle grupper.

Figur 11: Alle grupper - andeler av leddsetningstyper (per alle setninger)



Nominale leddsetninger ble brukt nesten like mye i alle tre grupper (figur 11) – andelen er litt høyere for gruppe A (9,4% per alle setninger), men gruppe B (8,5%) står på samme nivå med gruppe C (8,4%). Når vi ser på underkategoriene for nominale leddsetninger (tabell 11) så viser det seg at nominale leddsetninger som objekt ble betydelig hyppigere brukt av gruppe A (56% av alle leddsetninger) enn av gruppe B (33%) og C (30%). Nominale leddsetninger som subjekt forekom ikke i gruppe A, fantes bare 1 gang i gruppe B (0,4%) og frekvensen var også veldig lav for gruppe C (1%).

Tabell 11: Gruppe A - Andelen av underkategorier av leddsetninger (per alle leddsetninger)

	Gruppe A	Gruppe B	Gruppe C
Nominale leddsetninger...	56%	34%	31%
- som objekt	56%	33%	30%
- som subjekt	0%	0,4%	1%
Adjektiviske leddsetninger...	26%	32%	35%
- ...med <i>som</i> som objekt	0%	6%	6%
- ...med <i>som</i> som subjekt	19%	22%	23%
- ..som tilhører et tidsuttrykk	5%	2%	1%
- ...som tilhører et stedsuttrykk	2%	4%	5%
Adverbiale leddsetninger	22%	34%	34%

De største forskjellene som framgår fra tabell 11 forekommer dermed i bruk av de andre to underkategorier av leddsetninger. Hvis vi ser på andeler per alle setninger, har gruppe A prosentvist bare halvt så mange adjektiviske (A - 5,3%; B – 9%; C – 9,8%) og adverbiale

leddsetninger (A – 3,9%; B – 9%; C – 9%) per alle setninger som de andre grupper. Derimot står gruppe B og C igjen veldig nære hverandre.

Når vi ser på underkategorier av adjektiviske leddsetninger, så viser det seg at tekster i viderekommende gruppen likner igjen i høy grad morsmålstekstene. Adjektiviske leddsetninger med *som* som objekt utgjør 6% av alle leddsetninger i begge grupper, adjektiviske leddsetninger med *som* som subjekt representerer 22% av leddsetninger i gruppe B og 23% i gruppe C, adjektiviske leddsetninger som tilhører et stedsuttrykk brukes også like mye (B – 4%; C – 5%) og de som tilhører tidsuttrykk forekommer ganske sjeldent i begge grupper (B – 2%; C – 1%). Nybegynnergruppen skiller seg ut i bruk av disse strukturer. Adjektiviske leddsetninger med *som* som objekt fantes ikke i det hele tatt, mens setninger der *som* fungerer som subjekt (19% av alle leddsetninger) og setninger som tilhører et stedsuttrykk (2%) var mindre brukt enn i de andre grupper. Samtidig brukte gruppe A betydelig flere adjektiviske leddsetninger som tilhører et tidsuttrykk (5%).

Dette kan kanskje forklares med at adjektiviske leddsetninger som tilhører et tidsuttrykk, ofte danner en del av adverbialet i setningen. Jeg la merke til at der gruppe A har brukt en adjektivisk leddsetning tilknyttet et tidsuttrykk, har de andre brukt enten bare et adverbial i form av en enkel frase eller en adverbial leddsetning. Kvantitative resultater synes å støtte dette, siden gruppe A har en mye lavere andel av adverbiale leddsetninger (22% av alle leddsetninger) enn de andre grupper (34% for begge). Følgende eksempler som beskriver samme situasjon i fortellingen illustrerer den tendensen godt:

Eksempel 66: *En #da# natt, når gutten sov, #giga# gikk frosken ut av glasse#n#t. Om morgenen da gutten stå opp, så han at frosken hadde forsvant.* (student A5)

Eksempel 67: *Men, da gutten og hunden sin sovde, gikk frosken borte og i morgenen var glasset den hadde vært i tomt.* (student B2)

Eksempel 68: *Når kvelden kommer frister likevel sengen, og Per og Trofast legger seg, som de alltid gjør, sammen. Da de våknet neste morgen, savnet de frosken.* (informant C4)

Student A5 (eks 66) har beskrevet begge tidspunkter (kveld og morgen) med adjektiviske leddsetninger som er knyttet til tidsuttrykket og som sammen med dette danner adverbialet i oversetningen, samtidig som student B2 (eks67) og informant C4 (eks 68) har brukt en adverbial leddsetning i samme kontekst.

Bruk av formelt subjekt er også på ganske likt nivå i alle tre grupper – størrelsesorden er den samme, mens den relative forekomsten for gruppe A (5,1%) er igjen litt lavere enn for gruppe B (6,9%) og C (6,8%). Andelen av T-enheter som inneholder inversjon er også omtrent den samme, forskjellen mellom gruppe A (20,9%) og B (24,3%) er litt større, mens gruppene B og C (24,9%) har brukt like mye inversjon. Her må jeg samtidig igjen understreke at resultatene for inversjon inneholder også leddsetninger der inversjon ble brukt, selv om det ikke skulle ha blitt. Også for formelt subjekt har jeg tatt med setninger der denne strukturen skulle ha vært brukt, men ikke var det. Sammenlikning basert på korrekthet kommer jeg til å ta opp nedenfor.

4.3.3 Analysen av de lengste syntaktiske enheter

På grunn av det begrensede omfanget av oppgaven var det ikke mulig å ta med andre kvantitative målestokk for setningsintern kompleksitet enn bare *ord per setning*: jo lengre er setninger, jo mer tyder det på at setninger er komplekse. I teorikapittelet kommenterte jeg også relevansen av den målestokken og hvordan den er forskjellig fra andre målestokk som viser lengden av produksjonsenheter nettopp derfor at den bare handler om setningsintern kompleksitet (Norris og Ortega 2009:561). Ved den grundige gjennomlesningen av tekstene la jeg merke til de forskjellige syntaktiske trekk som er med til å skape setningsintern kompleksitet. Jeg vil illustrere dette aspektet ved et par næranalyser, mens problematikken må tas opp grundigere i en annen oppgave.

For å forstå bedre hvilke strukturer som spiller inn ved lengden av de forskjellige syntaktiske enheter, har jeg foretatt en næranalyse av de lengste ytringene, T-enhetene og setningene i hver gruppe (se tabell 12). Den lengste ytringen og T-enheten er omtrent like lange i alle tre grupper, for gruppe A er de litt kortere (resp 32 og 20 ord) enn for de andre grupper, mens gruppe B (resp 39 og 29 ord) er på samme nivå som morsmålsgruppen (resp 37 og 27 ord). Samtidig er forskjellen stor når det gjelder den lengste setningen i tekstene, her er setningslengden nesten samme i gruppe A (14 ord) og B (16 ord), men tydelig lengre (24 ord) i gruppe C.

De lengste ytringene er alle syntaktisk komplekse og det viser seg at man har brukt ganske ensartede syntaktiske ressurser i dem. Den lengste ytringen i gruppe A (eks 69) består av 3

Eksempel 69: Der gutten h rer #en# frosken #kv# /som/ kvakker og han med hunden ser frosken med kona og barna hans og gutten ser at frosken er glad fordi han kan slappe av med familie hans. (student A8)

Eksempel 70: *Samtidig* gutten sa inn i hyllet i treet, falt bikuben hunden hadde spillt med ned og alle bier var etter han og gutten falt ned av treet for ut av hyllet i treet kom en ugle som overrasket ham. (student B2)

Eksempel 71: Nå begynte gutten å forstå hvorfor frosken hadde rømt, og nå når han så hvor lykkelig frosken var for å være ute i det fri, bestemte han seg for å ikke ta frosken med seg hjem igjen. (informant C2)

I alle tre ytringer (eks 69-71) brukes det både sideordning og underordning, til og med andregradsunderordning. Både A8 og C2 har begynt ytringen med hovedsetningen, mens B2 bruker den grammatisk mer komplekse varianten og begynner ytringen med en foranstilt leddsetning. Forskjellen med morsmålsteksten er at både A8 og B2 har brukt to sideordnede setninger i sin ytring, mens C2 bare har en sideordnet konstruksjon og den følges straks av en underordning. B2 og C2 har også plassert en leddsetning (som ved C2 har til og med en andregradsleddsetning tilknyttet) innskutt i oversetningen, noe som A8 ikke har gjort.

² HS = Hovedsetning; HS... ...HS = Det finnes en innskutt leddsetning innenfor hovedsetningen; LS = Leddsetning; LS1 = førstegradsunderordnet leddsetning; LS2 = andregradsunderordnet leddsetning; SS = Sideordnet setning.

2 setninger, den i gruppe B (eks 73) av 5 setninger og den i gruppe C (eks 74) av 4 setninger.

Eksempel 72: *I bak av et gamle treet, som hadde falt ned ved siden av vatnet, fant gutten og hunden frosken.* (student A5)

HS... _____, ...HS _____.
LS (som) _____.

Eksempel 73: *Når gutten står på et stort stein og roper etter frosken, kommer bak steinen en hjort som hjelper å bære gutten på seg, slik at de kan gå fortere.* (student B9)

LS (Når) _____ SS (og) _____, HS _____
LS1 (som) _____, LS2 (slik at) _____.

Eksempel 74: *Mens Ole titter ned i et mistenkelig hull i bakken prøver Tassen å hoppe opp mot vepsebolet, før han innser at dette ikke fører noe steds hen.* (informant C3)

LS (Mens) _____ HS _____, LS1 (før) _____
LS2 (at) _____.

Her har det vært student A5 som har brukt en innskutt leddsetning i hovedsetningen (eks 72). Ellers er T-enheten til A5 kanskje litt enklere enn den til B9 og C3. Både student B9 og informant C3 har begynt T-enheten med en foranstilt leddsetning, B9 har til og med en sideordning som en del av leddsetningen (hennes T-enhet er også lengre enn den til C3). Begge har også både første- og andregrads leddsetning knyttet til oversetningen.

De største forskjellene synes å vise seg i de lengste setninger, da den lengste setningen i morsmålstekstene er betydelig lengre enn den i studenttekstene. Den lengste setningen i gruppe A er en hovedsetning (eks 75), den i gruppe B er en sideordnet setning (eks 76) og den i gruppe C danner en ytring i seg selv (eks 77).

Eksempel 75: *De vil ikke ha #gu# stor gutten og hunden å være i nærheten til dem,* (student A2)

Eksempel 76: *men frosken var ikke så glad i å #lev# bo i et glass resten av sitt #s#/liv.* (student B8)

Eksempel 77: *I den lille froskefamilien var en av froskeungene klar for å flytte hjemmefra og legge ut på nye eventyr sammen med gutten og hunden.* (informant C1)

Når vi ser på antallet av setningsledd, så er setningene ganske like. Setningen til A2 er litt vanskelig å analysere siden den ikke er grammatisk korrekt (inneholder en infinitt underordning *å være i nærheten til dem* som egentlig ikke passer hit), men i prinsipp bruker hun 5 setningsledd (subjekt, verbal, setningsadverbial, objekt og adverbial). B8 bruker også 5 setningsledd (konjunksjonal, subjekt, verbal, setningsadverbial og predikativ) mens C1 bare bruker 4 setningsledd (adverbial, verbal, subjekt, predikativ). Forskjellen må dermed ligge innenfor setningsleddene, på frasenivå – i antall attributter brukt og også i frasestrukturen, hvor komplekse fraser er og hva slags hode-dependent forhold de har. Både setningen til B8 og C1 har et predikativ som er utvidet i form av en lengre frase (understreket i eks 76 og 77), begge fraser inneholder en infinitivsfrase (infinitt underordning, dobbel understrekning), som i sin tur har 2 adverbialer knyttet til seg (i fet skrift). I alle tre setninger er det brukt en lengre adverbial som inneholder en preposisjonsfrase (til A2 består adverbialet av 4 ord; til B8 av 3 og 4; til C1 av 3 og 5 ord). Så hvor ligger forskjellen egentlig hvis alle synes å bruke de samme syntaktiske midler? Den eneste forskjellen som jeg kan peke på er at C1 bruker flere attributter foran substantiver (*den lille froskefamilien, en av froskeungene, nye eventyr*), som gjør setningen lengre.

Et annet trekk som jeg la merke til ved morsmålstekstene som jeg ikke fant hos studentene (med et enkelt unntak – eks 81), var bruk av setningsfragmenter eller setninger uten verb, både som selvstendige setninger men også som en del av en oversetning (eks 78-80).

Eksempel 78: *Men ikke nok med det, fra buskene kommer flere mindre frosker til syne.* (informant C3)

Eksempel 79: *Med dette vinker Ole og Tassen farvel til froskefamilien og lover å komme på snarlig besøk, men dog uten glasskrukken eller andre fangeredskaper.* (informant C3)

Eksempel 80: *Ugler er, som nevnt, veldig kloke og skjønner nesten alt.* (informant C4)

Eksempel 81: *Han klatret steinen, ropte etter frosken og ble en gang til overrasket – denne gangen av en hjort som stodde i buskene ved #siden av# /bak/ steinen.* (student B2)

Bruk av setningsfragmenter kan også være en forklaring på hvorfor setningene i gruppe C er lengre enn for gruppe A og B. Som vi kan se fra eksempel 78 og 80, er noen av disse setningsfragmenter veldig vanlige uttrykk på norsk, særlig i muntlig stil. Selv om tekstene i gruppe A og B var også preget av muntlig stil og lød som muntlige fortellinger, fantes det

ikke slike uttrykk. Dette er sannsynligvis knyttet til den leksikalske utviklingen og vi kan dermed igjen se hvordan den syntaktiske kompleksiteten kan være tett knyttet til utviklingen av ordforrådet og de allmenne kommunikative ferdigheter.

Siden den ovenstående næranalysen bare så på enkelte eksempler fra tekstene i hver gruppe, er det ikke mulig å dra altfor vidtgående konklusjoner basert på denne. Allikevel synes jeg at eksemplene illustrerer godt tendensene som også avspeiler seg i den kvantitative analysen, samt viser det behovet for videre undersøkelser på mer detaljert setningsinternt nivå, som vi ikke har mulighet for å ta opp i rammene av denne avhandlingen.

Tabell 12: Den lengste ytringen, T-enheten og setningen i hver gruppe.

Den lengste ytringen	Den lengste T-enheten	Den lengste setningen
Student A8 – 32 ord: <i>Der gutten hører #en# frosken #kv# /som/ kvakker og han med hunden ser frosken med kona og barna hans og gutten ser at frosken er glad fordi han kan slappe av med familie hans.</i>	Student A5 – 20 ord: <i>I bak av et gamle treet, som hadde falt ned ved siden av vatnet, fant gutten og hunden frosken.</i>	Student A2 – 14 ord: <i>De vil ikke ha #gu# stor gutten og hunden å være i nærheten til dem,</i>
Student B2 – 39 ord: <i>Samtidig gutten sa inn i hyllet i treet, falt bikuben hunden hadde spillt med ned og alle bier var etter han og gutten falt ned av treet for ut av hyllet i treet kom en ugle som overrasket ham.</i>	Student B9 – 29 ord: <i>Når gutten står på et stort stein og roper etter frosken, kommer bak steinen en hjort som hjelper å bære gutten på seg, slik at de kan gå fortere.</i>	Student B8 – 16 ord: <i>men frosken var ikke så glad i å #lev# bo i et glass resten av sitt #s#/iv.</i>
Informant C2 – 37 ord: <i>Nå begynte gutten å forstå hvorfor frosken hadde rømt, og nå når han så hvor lykkelig frosken var for å være ute i det fri, bestemte han seg for å ikke ta frosken med seg hjem igjen.</i>	Informant C3 – 27 ord: <i>Mens Ole titter ned i et mistenkelig hull i bakken prøver Tassen å hoppe opp mot vepsebolet, før han innser at dette ikke fører noe steds hen.</i>	Informant C1 – 24 ord: <i>I den lille froskefamilien var en av froskeungene klar for å flytte hjemmefra og legge ut på nye eventyr sammen med gutten og hunden.</i>

4.3.4 Syntaktisk korrekthet

Til sist skal vi se på andelen av de syntaktisk korrekt formulerte strukturer (figur 12). Selv om jeg forventet å finne eksempler på avvik fra normen også i morsmålstekstene, la jeg ikke merke til noen. Således er korrekthetsprosenten for gruppe C 100% i alle kategorier. Dermed skal jeg i den delen av analysen sammenlikne bare gruppe A og B.

Hvis vi ser på et generelt plan, så er forskjellen i syntaktisk korrekthet forholdsvis stor mellom gruppene. Andelen av syntaktisk korrekte T-enheter ligger ganske høyt for gruppe B (95%), de har oppnådd nærmest førstespråksliknende kompetanse når det gjelder syntaktisk korrekthet. Resultatene for gruppe A ligger derimot lavere (83%) og de har problemer mer korrekthet i flere syntaktiske kategorier. Det finnes allikevel noen kategorier som begge gruppene har mestret på et høyt nivå: enkelytringer (A – 91%; B – 94%) og sideordnede setninger (A – 90%; B 98%). Samtidig ser vi at i mange andre kategorier er forskjeller betydelig større.

Gruppe A har mer problemer ved hovedsetninger (83% korrekte), mens gruppe B presterer veldig høyt også i den kategorien (98%). Ved nærmere analyse av tekstene, viser det seg at det forekommer 2 typer problemer ved hovedsetningene. De fleste ukorrekte hovedsetninger i gruppe A bryter mot V2 regelen ved bruk av inversjon, noe som ikke var tilfellet i gruppe B. Mange hovedsetninger i gruppe A begynte med et adverbial enten i form av et ord, en frase eller en adverbial leddsetning, men fortsatte med normal i stedet for invertert ordstilling, som i følgende eksempler:

Eksempel 82: *Etterpå gutten og hunden gikk hjem og de froskene kvakket sammen til kveld.* (student A1)

Eksempel 83: *Hvis gutten sover, frosken går ut på fiske#flaske#/glasse/ og går ut /for/ å se landet.* (student A8)

Det andre problemet som gjorde at korrekthetsprosenten for hovedsetninger ble lavere var noe som forekom i hovedsetninger for både gruppe A og B. Det var unnlatelse av formelt subjekt, særlig i setninger av typen «*Der var...*» i stedet for «*Det var...*» (dette ble nærmere beskrevet i kapittel 4.1.3), men også i andre typer setninger der man skulle ha brukt et formelt subjekt (eks 84-87).

Eksempel 84: *Nærhaten fjellet var en hjort og hjorten tok Nils og begynte å løpe.* (student A7)

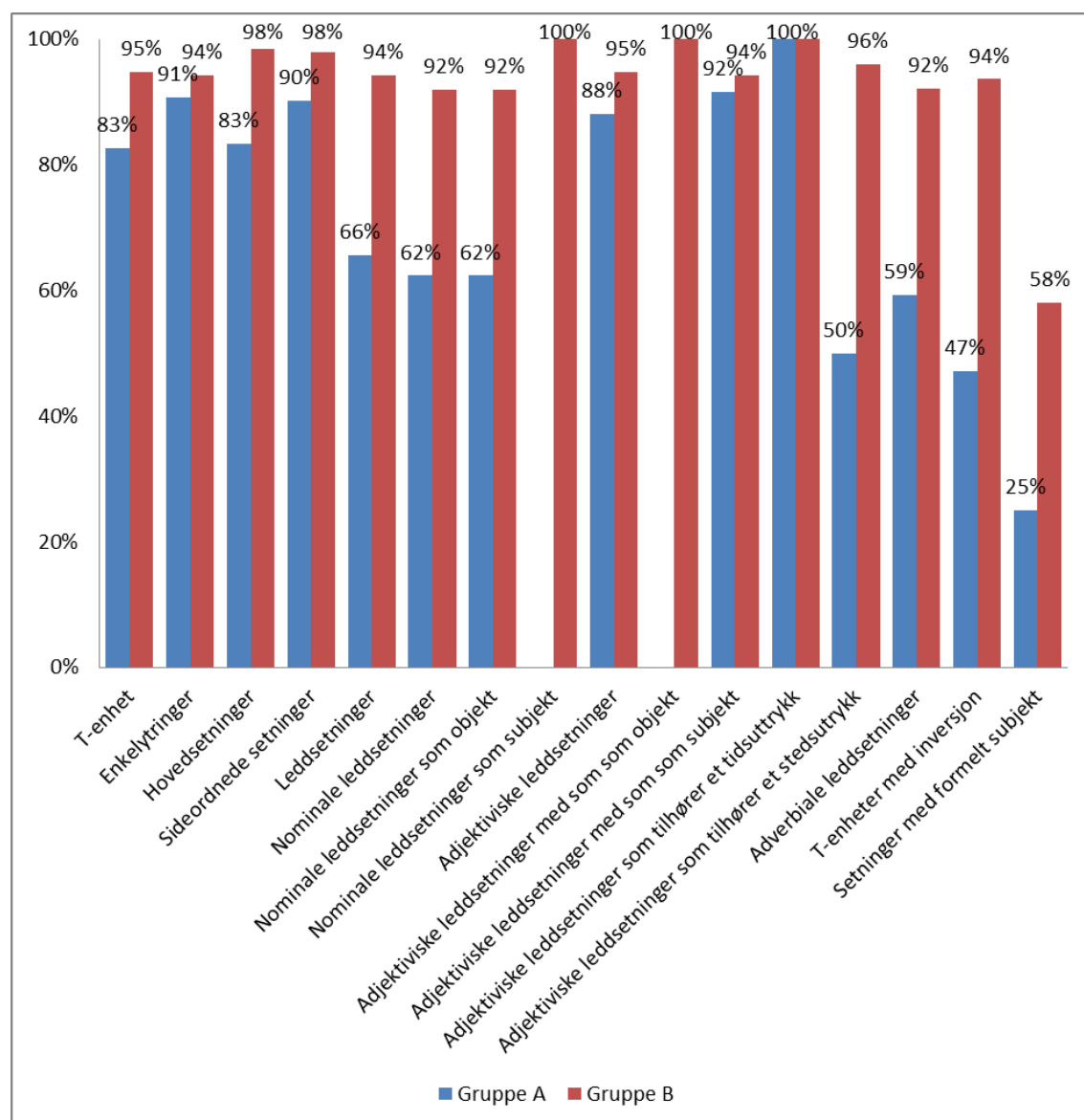
Eksempel 85: *Der vær også en hjørt som vær ikke så god.* (student A11)

Eksempel 86: *Der var bare en muldvarp som likte det ikke at noen ropte i hjemmet hans.* (student B2)

Eksempel 87: *Når uglen og bier har sluttet jakte kommer ## en hjort som tar Morten på hode sin.* (student B10)

De ovenstående eksemplene illustrerer også korrekthets-statistikken over setninger med formelt subjekt (figur 12). Bruk av formelt subjekt i setningene viser seg å være vanskelig for begge gruppene – særlig lav er korrekthetsprosenten for gruppe A – 25%, men også gruppe B klarer seg i bare halvdelen (58%) av setninger som har eller skulle ha formelt subjekt.

Figur 12: *Gruppe A og B - andelen av syntaktisk korrekte strukturer*



Inversjon derimot synes å være noe som gruppe B har allerede tilegnet seg godt (94% korrekte), mens gruppe A fremdeles har større problemer med å holde til V2 regelen i setningene der man har foranstilt et adverbial enten i form av en frase eller en leddsetning (bare 47% korrekte).

Det andre området hvor vi kan se store forskjeller, er leddsetninger. Gruppe B har ganske høy korrekthetsnivå i alle leddsetninger (94%). Over 90% av leddsetningene er korrekte i alle underkategoriene. Samtidig kan vi se at gruppe A har fremdeles en vesentlig mengde problemer med leddsetningene (bare 66% korrekte). De mest problematiske underkategorier her er nominale leddsetninger som objekt (62% korrekt), adjektiviske leddsetninger som tilhører stedsuttrykk (50 % korrekt) og adverbiale leddsetninger (59% korrekt). Problemene som studenter i gruppe A hadde i forbindelse med leddsetningene har jeg beskrevet i detalj i kapittel 4.1.3.

Alt i alt kan vi si at begge studentgrupper bruker komplekse syntaktiske strukturer. Gruppe A har naturligvis litt færre komplekse strukturer, særlig på de høyere kompleksitetsnivåer (som leddsetninger). Allikevel kan vi finne en hel del eksempler på høy grad av kompleksitet allerede ved begynnergruppen. Gruppe A skiller seg mest ut fra de andre grupper med sitt lavere korrekthetsprosent i de fleste syntaktiske kategorier. Tekstene i gruppe B likner allerede i stor grad morsmålstekstene når det gjelder valg av syntaktiske strukturer og kompleksitet. Gruppe B skårer også høyt ved korrekthet i de fleste analysekategoriene. Det som mest skiller syntaksen i studentenes tekster fra morsmålstekstene er den setningsinterne kompleksiteten, som speiles av mye lengre setninger i gruppe C.

5 Diskusjon og konklusjon

Vi har nå ved hjelp av både kvantitative målestokk og kvalitativ næranalyse sett på hva som kjennetegner innlærersyntaksen ved to grupper av voksne innlærere av norsk, samt i hvilke trekk den er lik eller forskjellig fra målspråkets syntaks. Sammenlikningen mellom gruppene viste ganske mange likheter – likheter i valg av de forskjellige syntaktiske strukturer mellom alle tre grupper og samsvar mellom kategoriene der begge studentgruppene hadde problemer med. Når det gjelder syntaktisk korrekthet så finnes det kategorier der begge studentgruppene står overfor utfordringer – for gruppe A finnes det, logisk nok, flere enn for gruppe B.

Hva som kjennetegner syntaksen på de respektive utviklingsnivåene og i hvilke trekk er det forskjellig fra språkbruket til morsmålstalere? Nybegynnergruppen lager kortere syntaktiske enheter enn de andre grupper, samtidig som sideordnings- og underordningsgraden står ganske nært til de andre. Det finnes litt flere enkelytringer og sideordnede setninger og litt mindre underordning, samt forekommer det vanligvis bare førstegradsunderordning. Når det gjelder forskjellige typer leddsetninger, kan man se en sterk dominans av nominale leddsetninger hos gruppe A. De andre typer leddsetninger er mindre brukt enn hos de andre gruppene. Det som skiller nybegynnere mest fra både gruppe B og C, er lavere antall syntaktisk korrekte strukturer.

Den viderekommende gruppen likner allerede i de fleste analysekategoriene med morsmålstalere. De bruker litt flere enkelytringer, men sideordnings- og underordningsgraden er stort sett samme som i gruppe C. Dette gjelder også valg av leddsetningstypene. De største forskjellene viser seg i lengden av ytringer, T-enheter og setninger som er mindre i den viderekommende gruppen enn i morsmålsgruppen. Samt klarer studentene ikke å forme alle syntaktiske strukturer helt korrekt ennå.

Hvorvidt brukes komplekse syntaktiske strukturer på de forskjellige utviklingsnivåene? Gjennom hvilke trekk kommer den syntaktiske kompleksiteten til syne? Hvilke typer av strukturer studentene bruker? Komplekse strukturer er i bruk ved begge studentgrupper.

Den syntaktiske kompleksiteten viser seg i gruppe A mest gjennom sideordning, litt mindre gjennom underordning og ved forsøk å bruke inversjon. Når det gjelder leddsetningstyper, har vi eksempler på nesten alle typer leddsetninger (selv om noen typer er representert av bare noen få eksempler). Gruppe B bruker mer underordning og inversjon, samt er alle leddsetningstyper representert i deres tekster. Bruk av formelt subjekt er også flittigere og mer framgangsrik i gruppe B.

Hvordan kan likhetene og forskjellene mellom grupper forklares? Selfølgelig er en av årsakene til forskjellene det at gruppe A ikke har fått undervisning i de mer komplekse strukturer ennå, men den syntaktiske kompleksiteten er ikke avhengig av bare det. Ved mange eksempler i tekstene, kunne vi se hvordan utviklingen av leksikalske og syntaktiske kunnskaper er avhengige av hverandre. I mange tilfeller ved gruppe A kan vi se at problemene med syntaks kan spores tilbake til manglende leksikalske kunnskaper. Samtidig ser vi fra materialet i Gruppe B hvordan omfattende leksikalske kunnskaper (for eksempel et bredt utvalg av subjunksjoner) medfører hyppigere bruk av komplekse syntaktiske strukturer. Dette tar oss tilbake til Pallotti (2009: 593) sin diskusjon om avhengigheten mellom de forskjellige språklige dimensjoner og at de egentlig ikke står adskilt. Videre undersøkelser skulle være nødvendige for å se på hvordan den syntaktiske kompleksiteten egentlig utvikler seg i forhold til den syntaktiske.

Mange likhetene mellom gruppe A og B, særlig når det gjelder syntaktiske strukturer som de sliter med, kan forklares med at morsmålet estisk ikke bruker slike strukturer eller har forskjellige syntaktiske regler enn norsk. Vi kan også ved noen tilfeller betrakte påvirkning av andre fremmedspråk som studentene har lært seg før – for eksempel engelsk.

Jeg synes at ved siden av forskjellene og utviklingsrelaterte særpreg er det også viktig å forstå likhetene. Pallotti (2009) argumenterer hvordan forskere hele tiden prøver å finne på bedre og nøyaktigere metoder og målestokk for å tydeligst vise forskjellene mellom forskningsobjektene. Samtidig mener han at «en målestokk kan være vitenskapelig gyldig og informativ, selv om det ikke viser noen forskjell mellom grupper eller individer man forsker» (Pallotti 2009: 590). Det er naturlig at forskning er opptatt av variasjoner og forskjeller, men det er også viktig å legge merke til konstanter og likheter (Pallotti 2009:

590). Jeg er enig med Pallotti og derfor synes jeg at det kan være nyttig å konsentrere seg herved på det som er likt i studentenes mellomspråk på begge nivåer.

For det første er det viktig å forstå at voksne innlærere (universitetsstudenter) har et høyt nivå av språkkompetanse i sitt eget morsmål, i tillegg har studentene i både gruppe A og B mestret mer enn ett eller to fremmedspråk. Dette skaper grunnlaget til hvilket utgangspunkt de har når de begynner å lære et nytt språk. Den intuitive antakelsen, at nybegynnere bare bruker enkelytringer, holder ikke. I hvert fall når språket som tilegnes er det tredje, fjerde eller femte fremmedspråket. Vi kan tydelig se fra materialet og analysene i denne oppgaven, at behovet og viljen at bruke mer komplekse syntaktiske strukturer, finnes både på begynnernivået og selvfølgelig også hos viderekommende studenter. I situasjonen der studentenes egne språklige ressurser eller kunnskaper i norsk ikke strekker til, har de andre språk og ikke minst morsmålet som et forbilde som de kan støtte på. Hvis studenten er vant til å uttrykke mer komplekse språklige forhold på andre språk, så er det nokså sannsynlig at hun vil gjøre det på norsk også. Derfor synes jeg at det er et aspekt, som man bør ta hensyn til ved undervisningen også. Selvfølgelig må man konsentrere seg på det grunnleggende stoffet i begynnelsen av språktilegnelsen, men man skal ikke glemme at studentene behøver veiledning i de komplekse strukturene også - de prøver seg på dem allikevel.

Hvilke av strukturene har studentene tilegnet framgangsrikt og hvilke ikke? Hvilke tendenser finnes det med hensyn til strukturer som studentene har problemer med? Den syntaktiske korrektheten, som sagt, var lavest i begynnergruppen, og ganske høy i den viderekommende gruppen. Nybegynnere har mindre problemer i enkelytringer og sideordning og flere problemer i leddsetninger. Gruppe B har tilegnet seg de fleste kategoriene ganske bra, det finnes en tendens mot litt flere ukorrekte setninger i de forskjellige leddsetningstypene.

Vedrørende den syntaktiske korrektheten av de forskjellige strukturene dreier det allikevel ikke så mye om de konkrete setningstypene, enn om de to syntaktiske konstruksjoner som kan brukes i forskjellige setninger: formelt subjekt og inversjon. I begge gruppene var det få studenter som kunne lage setninger med formelt subjekt slik som det var naturlig hos morsmålsbrukere, da tekstenes natur og innhold forutsetter bruk av presenterende setninger. Dette kan selvfølgelig også være avhengig av hvilken type oppgave vi har å

gjøre med, kanskje er det at man kommer inn med nye karakter og må presentere nye scener noe som spesielt karakteriserer fortellinger. Samtidig er det sannsynlig at bruk av formelt subjekt er problematisk for estere også ved andre typer av tekster, siden et slikt syntaktisk fenomen ikke eksisterer i estisk.

Samtidig som bruk av formelt subjekt er noe som studentene enten ikke klarte å lage korrekt eller unngikk helt, er inversjon en av slike kategorier som de prøvde på veldig flittig, men (særlig på begynnernivået) ikke klarte korrekt. Særlig på begynnernivået forekom det brudd mot V2 regelen, samt forsøk på å bruke inversjon i leddsetninger (som ikke kan gjøres på norsk). Her er det igjen sannsynlig, at når en voksen innlærer forteller en historie, er det ikke naturlig for ham/henne å begynne alle setninger med subjekt. Man prøver å gjøre sin fortelling mer spennende og dynamisk, legge mer vekt på forskjellige aspekter i setningen og strukturere informasjonen på forskjellige måter, dermed har man behov for å flytte et annet setningsledd fram i setningen.

Hvis hensikten med andrespråkstilegnelsen / undervisningen er å hjelpe studenter å oppnå morsmålsliknende kompetanse, er det viktig å se på hva de klarer på lik linje med morsmålstalende informanter og hva er det som er mangelfullt ved språket deres. Konklusjonen, som jeg kan trekke på bakgrunn av min analyse, er at ved valg av forskjellige setningstyper – bruk av enkelytringer, sideordnings- og underordningsgraden – skiller studentenes tekster ikke mye fra morsmålstekstene. Samtidig som vi kan se et klart forskjell i setningslengden, som øker seg fra gruppe A til B til C. Forskjellen synes dermed å ligge i den syntaktiske kompleksiteten innenfor setningen – setningsstrukturen, frasestrukturen, bruk av attributter osv. Dette passer godt inn i sammenhengen av Norris og Ortega sine diskusjoner (2009: 562-563) om at vi må utvide vår forståelse av syntaktisk kompleksitet og at utviklingen av syntaksen ikke bare dreier seg om «mer underordning» eller «flere og flere setninger per T-enhet» men også om det som skjer setningsinternt.

Det som det eventuelt må legges mer vekt på i undervisningen er da ikke nødvendigvis hvordan man plasserer setningsadverbialer i leddsetningen, men heller hvordan man danner mer komplekse fraser. Grundigere analyse av kompleksiteten innenfor setninger og syntaksens forbindelse med den leksikalske utviklingen vil være nødvendig i framtiden – både på det samme materialet men også ved andre liknende forskningsprosjekter. Jeg vil

anbefale å bruke materialet som ble samlet inn i forbindelse med dette prosjekt for framtidige andrespråksundersøkelser, siden materialet inneholder mange interessante momenter og kan analyseres fra flere forskjellige synspunkter.

6 Eestikeelne resümee

Tartu Ülikool

Germaani, romaani ja slaavi filoloogia instituut

Skandinavistika osakond

Norra keel ja kirjandus

Töö pealkiri: "Konn, kus sa oled?" – süntaktiline kompleksus eesti norra keele tudengite kirjalikes jutustustes.

Magistritöö

Autor: Karoliine Piksarv

Juhendaja: Anu Laanemets, PhD

Käesoleva magistritöö eesmärgiks on kaardistada norra keele õppijakeele süntaksi kompleksusastet ja variatiivsust täiskasvanud keeleõppuritel, kelle emakeeleks on eesti keel. Uurimus põhineb tudengite kirjalikul produktsioonil ning vaatleb süntaktilist kompleksust kahe tasemegrupi – algajate ning kesktasemel keeleõppurite – tekstides. Algajate grupp (n=11) oli materjali kogumise hetkel õppinud norra keelt valikainena umbes poole aasta jooksul, kesktaseme grupi tudengid (n=10) õppisid norra keelt aga põhierialana ning olid läbinud 4 semestri jooksul erinevaid norrakeelseid loengukursusi ja seminare. Töös võrreldakse õppijakeele süntaksit ka norra keelt emakeelena kõnelevate täiskasvanud informantide (n=5) lauseehitusega.

Analüüsi tulemusena saame paremini aru süntaktilise kompleksuse arengust teise keele omandamise kontekstis. Samuti loodab töö autor oma arutlustega anda panuse norra keele õpetamist puudutavasse diskussiooni, et tulevikus saaksime oma tudengite arengut toetada parimal võimalikul viisil.

Uuringu käigus leiti vastused järgnevatele küsimustele:

- 1) Mis iseloomustab nende kahe tasemegrupi tudengite süntaksit ning mille poolest erineb see emakeelekõnelejate omast?
- 2) Mil määral kasutatakse kompleksseid süntaktilisi struktuure erinevatel arengutasemetel? Milliste süntaktiliste nähtuste kaudu avaldub kompleksus tudengite tekstides? Milliseid struktuure tudengid kasutavad?

- 3) Millised süntaktilised struktuurid on omandatud edukalt ja millised mitte? Milliseid tendentse võime märgata struktuuride osas, millega keeleomandajatel on probleeme?
- 4) Millega saab seletada sarnasusi ja erinevusi gruppide vahel?

Süntaktilist kompleksust on antud töös analüüsitud lause tasandil, analüüs ei hõlma lauseüleseid kategooriaid. Samuti on magistritöö piiratud mahu tõttu jäänud käsitlusest kõrvale osalausest väiksemad analüüsiühikud – fraasi ja lauseliikmete siseseid kompleksuse ilminguid on veidi puudutatud vaid kvalitatiivses analüüsis.

Uurimus põhineb tudengite kirjalikel jutustustel, mis tuginevad laste pildiraamatu ”*Frog, where are you?*” piltidele (Mayer 1969). Antud pildiseeria on leidnud keeleteaduslikes uurimustes laialdast kasutust ning selle abil on uuritud nii lapsekeele (nt. Berman ja Slobin 1994), varajase kakskeelsuse (nt. Bennet-Kastor 2002) kui ka võõrkeelte (nt. Tortani 2012) arengut.

Töö teoreetiline vaatepunkt paigutub rakenduslingvistikas laialdast kasutust leidvate CAF (*Complexity, Accuracy, Fluency*) uuringute alla ning vaatleb süntaksi arengut kompleksuse näitajate kaudu. Süntaktiline kompleksus on siinkohal defineeritud kui keelekasutuses väljenduvate erinevate süntaktiliste struktuuride repertuaar / variatsioon, samuti nende struktuuride keerukusaste (Ortega 2003:492). Süntaktiliste struktuuride varieeruvust ning keerukust kirjeldatakse nii kvantitatiivse kui kvalitatiivse analüüsi abil. Kasutusel on mitmed uurimisvaldkonnas laialt levinud mõõdikud (Norris ja Ortega 2009:561-562), mille kaudu saame tunnistust eri kompleksusastmete nähtustest:

- Üldine kompleksus, mida mõõdetakse erinevate süntaktiliste ühikute pikkusega.
- Kompleksus, mis väljendub koordineerimise tasemes (kui palju esineb rindlauseid).
- Kompleksus, mis väljendub subordinatsiooni tasemes (kui palju esineb põimlauseid).
- Osalauseisisene kompleksus, mis väljendub fraasi ja lauseliikmete tasandil.
- Kompleksus kui struktuuraalne varieeruvus, st. mitmekesiste süntaktiliste struktuuride kasutamine.

Süntaktilise komplekssuse määramiseks on kõiki tekste analüüsitud järgmisi mõõdikuid kasutades:

- lausete keskmine pikkus (sõnade arv) – üldise komplekssuse näitaja
- T-ühikute keskmine pikkus (sõnade arv) – üldise komplekssuse näitaja
- osalausete keskmine pikkus (sõnade arv) – üldise komplekssuse näitaja
- keskmine T-ühikute arv lause kohta – näitab rindlausete osakaalu
- keskmine osalausete arv T-ühiku kohta – näitab põimlausete osakaalu
- keskmine kõrvallausete arv osalausete kohta – näitab mitmenda astme kõrvallausete osakaalu

Selleks, et analüüsida erinevate süntaktiliste struktuuride kasutussagedust tekstides, on kõik osalaused märgendatud järgnevasse kategooriatesse:

- nominaalsed kõrvallaused;
 - nominaalsed kõrvallaused pealause sihitisena;
 - nominaalsed kõrvallaused pealause alusena;
- täiendlaused ehk adjektiivsed kõrvallaused;
 - täiendlaused, kus *som* on sihitiseks;
 - täiendlaused, kus *som* on aluseks;
 - täiendlaused, mis kuuluvad ajaväljendi juurde;
 - täiendlaused, mis kuuluvad kohaväljendi juurde;
- määruslaused ehk adverbiaalsed kõrvallaused;
- T-ühikud, mis sisaldavad inversiooni;
- Osalaused, kus on kasutatud formaalset alust.

Samuti on iga osalause puhul määratud, kas see on süntaktiliselt korrektne või mitte ning arvutatud välja korrektsete struktuuride osakaal kõigis ülalpoolnimetatud kategooriates.

Analüüsi tulemustena leiti, et erinevate süntaktiliste struktuuride esinemissagedustes ei olnud gruppide vahel märkimisväärsed erinevusi – suurusjärgult olid tulemused sarnased nii gruppidel omavahel kui võrreldes tudengite tekste emakeelekõnelejadega (vt. ka graafikuid 9, 10, 11 ja tabel 11 peatükis 4.3.1). Süntaktiliselt keerukad struktuurid leidsid

ühtjagu kasutust mõlemas grupis, kusjuures edasijõudnute grupis leidis näiteid kõigist analüüsikategooriatest ning algajatel jäid puudu vaid nominaalsed kõrvallauseid pealause alusena ning täiendlauseid, kus *som* on sihitiseks (tabel 11). Algajate grupis kasutati protsentuaalselt pisut enam liht- ja rindlauseid, põimlause osakaal oli veidi väiksem kui edasijõudnutel (graafik 10). Algajatel esines enam nominaalseid kõrvallauseid, samas kui täiendlauseid ning määruslauseid osakaal oli teistest gruppidest madalam (tabel 11). Samuti kasutasid algajad veidi vähem inversiooni. Edasijõudnute süntaktiliste struktuuride valik oli aga üsna sarnane emakeelekõnelejadega, pea kõigis kategooriates olid tulemused emakeelegrupiga samal tasemel.

Suurimad erinevused ilmnest üldistes kompleksuse näitajates (graafik 9). Läbi kolme grupi suureneb ühtlaselt keskmine sõnade arv lauses, T-ühikus ning osalauses, samas kui rind- ja põimlause osakaal kõigi osalausede suhtes jääb suhteliselt samale tasemele. Sellest järeldub, et süntaktilised struktuurid peavad pikenema läbi arengutasemete millegi muu arvelt – miski peab muutuma osalause siseselt. Kvalitatiivse lähianalüüsi tulemusena on näha, et edasijõudnute aga eriti emakeelekõnelejad osalaused on pikemad ning sisaldavad enam täiendeid, infinitiivfraase ning emakeelekõnelejad puhul ka lausefragmente. Leidis ka mitmeid näiteid üsna keeruka struktuuriga ja pikkadest lihtlausestest.

Teiseks oli näha suuri erinevusi gruppide vahel süntaktilises korrektsuses (vt. ka graafik 12, peatükis 4.3.4). Algajate grupi tulemused olid korrektsuse osas ootuspäraselt kõige madalamad. Kõigist T-ühikutest olid korrektsed 83%. Siiski leidis ka algajatel kategooriaid, kus süntaktiliselt õigete lause osakaal on üsna kõrge: lihtlausestest on õiged 91%, rindlausestest 90%, täiendlauseid, kus *som* on aluseks 92% ning ajamääruse juurde kuuluvad täiendlauseid 100%. Kahes viimases kategoorias leidis lauseid küll väga vähe, kuid tähelepanuväärne on seegi, et algtasemel keeleõppijate seas leidis neid, kes oskavad selliseid struktuure veatult koostada.

Algajate grupis olid kõige tagasihoidlikumad tulemused korrektsuse osas kõrvallausestes (vaid 66% õiged). Enim tehakse vigu lauseadverbiaali asukohaga kõrvallausestes, samuti tekitab ebakorrektsed lauseid püüd kasutada kõrvallause inversiooni ehk asetada määrus kõrvallause algusse. See pole norra keeles kõrvallause kindla sõnajärje tõttu võimalik, küll

aga võib siin näha emakeele mõju, kuivõrd eesti keeles kõlab määrusega alustatud kõrvallause täiesti tavapäraselt. Inversioon ehk alusest erineva lauseliikme paigutamine lause algusse on tekitanud probleeme ka lihtlausetes, pealausetes ning rindlausetes. Vaid 47% inversiooni sisaldavatest lausetest on korrektselt moodustatud – peamiselt ei ole peetud kinni reeglist asetada tegusõna alati lauses teisele kohale. Oluliseks probleemkohaks on algajate grupis ka formaalse aluse kasutamine. Lausetest, mis sisaldasid või oleks pidanud sisaldama formaalset subjekti, on õiged vaid 25%.

Edasijõudnute grupi süntaktilise korrektsuse tulemused on aga üsna kõrgel tasemel ning nii mõneski kategoorias lähenevad emakeelekõnelejatele. Kõigist T-ühikutest olid korrektsed lausa 95%. Kõigis kategooriates, välja arvatud üks (formaalne subjekt), on tulemus üle 90%. Taaskord leidis vigu kõrvallausetes, kusjuures ka edasijõudnute seas leidis näiteid inversiooni kasutamisest kõrvallauses (mis norra keeles pole süntaktiliselt korrektne). Kõige problemaatilisem kategooria edasijõudnute grupis oli laused, mis sisaldasid või oleksid pidanud sisaldama formaalset alust (vaid 58% korrektseid). Nii algajate kui edasijõudnute puhul oli märgata esitluslausetes puhul (kus norra keeles peaks kasutama formaalset alust) inglise või eesti keele mõju. Näiteks oli alustatud lauset sageli, mitte formaalse aluse, vaid määrusega: „*There is / are*“ või „*Seal oli / on...*“, mis norra keeles kõlab aga väga ebaloomulikult.

Uurimustulemustest nähtub, et täiskasvanud keeleõppijad (ülikooli tudengid), kel on kõrgtasemel emakeeleoskus ning kes valdavad mitmeid võõrkeeli, on võimalus toetuda neile oskustele uute keelte omandamisel. Seetõttu ei ole neile omane piirduda isegi algtasemel vaid lihtsate süntaktiliste struktuuridega ega kasutada ainult lihtlauseid. Kasutades teisi keeli lisaressursina üritavad nad moodustada keerukaid konstruktsioone, ka siis kui nad ei ole neid tegelikult õppinud. Seetõttu oleks otstarbekas tutvuda keerulisemate konstruktsioonidega õppetöös juba algtasemel.

Kui keeleõppe eesmärgiks on saavutada võimalikult emakeelekõnelejatele sarnane kompetents, siis vajaksid õpilased õppetöö raames rohkem tuge inversiooni ning formaalse aluse omandamisel. Samuti tuleks keskenduda nii algajate kui edasijõudnute tasemel lisaks keerukatele lausekonstruktsioonidele ka osalausesisestele struktuuridele ning

fraasiõpetusele. Oluliseks abiks oleks siinkohal edaspidised süntaktilise kompleksuse uuringud, mis keskenduksid eelkõige osalause sisestele nähtustele.

Litteraturliste

- Bardovi-Harlig, Kathleen. 1992. A Second Look at T-Unit Analysis: Reconsidering the Sentence. *TESOL Quarterly* Vol. 26, No. 2.
- Bennet-Castor, Tina. 2002. The “frog story” narratives of Irish–English bilinguals. *Bilingualism: Language and Cognition* Volume 5 / Issue 02 pp. 131-146
- Berman, R., Slobin, D. I. 1994. Relating events in narrative: A crosslinguistic developmental study. Hillsdale, NJ: Lawrence Erlbaum.
- Ellingsen, E., Mac Donald, K. 2004. *På vei. Tekstbok. Norsk og samfunnskunnskap for voksne innvandrere*. Cappelen.
- Ellis, R. 2003. Task-based Language Learning and Teaching. Oxford University Press
- Faarlund, J. T., Lie, S., Vannebo, K.I. 1997. *Norsk Referansegrammatikk*. Universitetsforlaget. Oslo
- Gaies, Stephen J. 1980. T-Unit Analysis in Second Language Research: Applications, Problems and Limitations. *TESOL Quarterly*, Vol. 14, No. 1.
- Golden, A., Mac Donald K., Ryen, E. 2011. *Norsk som fremmedspråk. Grammatikk*. Universitetsforlaget. Oslo
- Hagen, J. E. 1998. Norsk grammatikk for andrespråklærere. Oslo: Ad Notam.
- Housen, A., Kuiken, F. 2009. Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics* 30/4. pp. 461-473. Oxford University Press
- Jiang, Wenyang. 2013. Measurements of Development in L2 Written Production: The Case of L2 Chinese. *Applied Linguistics* 34-1. pp. 1-24. Oxford University Press.
- Johansen, Hilde. 2008. Inversjon i norsk innlærerspråk. En undersøkelse av variasjonsmønstre i skrevne tekster. *NOA norsk som andrespråk. Årgang 24. 2/2008*
- Larsen-Freeman, Diane. 1978. An ESL Index of Development. *TESOL Quarterly* Vol. 12, No. 4. pp. 439-448

Larsen-Freeman, Diane. 2009. Adjusting Expectations: The Study of Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics* 30/4. pp. 579-589

Lennon, P. 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning* 40. pp. 387-417

Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Press

Norris, J. M., Ortega, L. 2009. Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics* 24-4. pp. 555-578

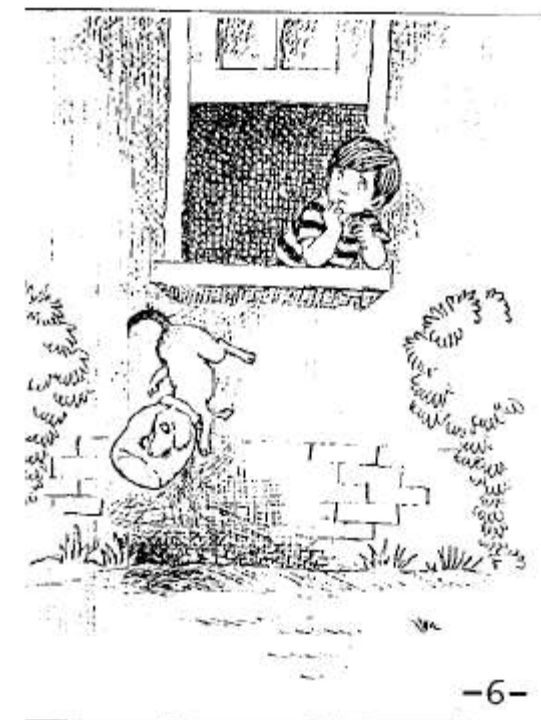
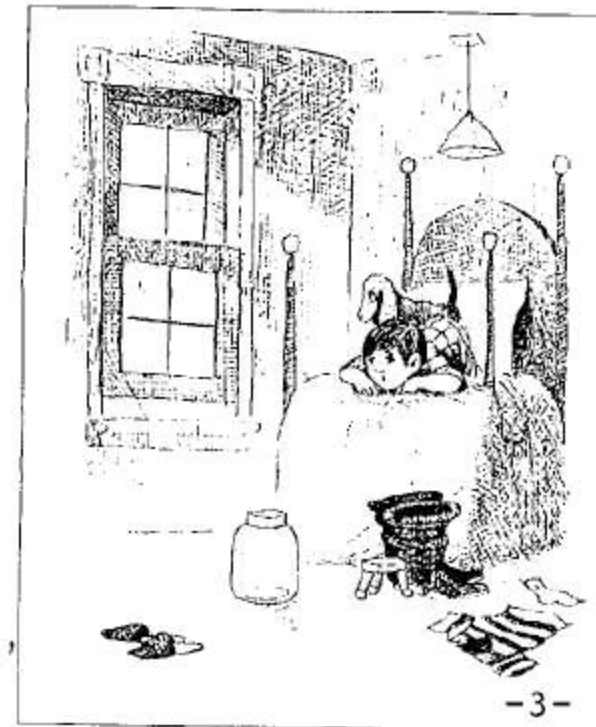
Ortega, Lourdes. 2003. Syntactic Complexity Measures and their Relationship to L2 Proficiency: A Research Synthesis of College-level L2 Writing. *Applied Linguistics* 24-4. pp. 492-518. Oxford University Press.

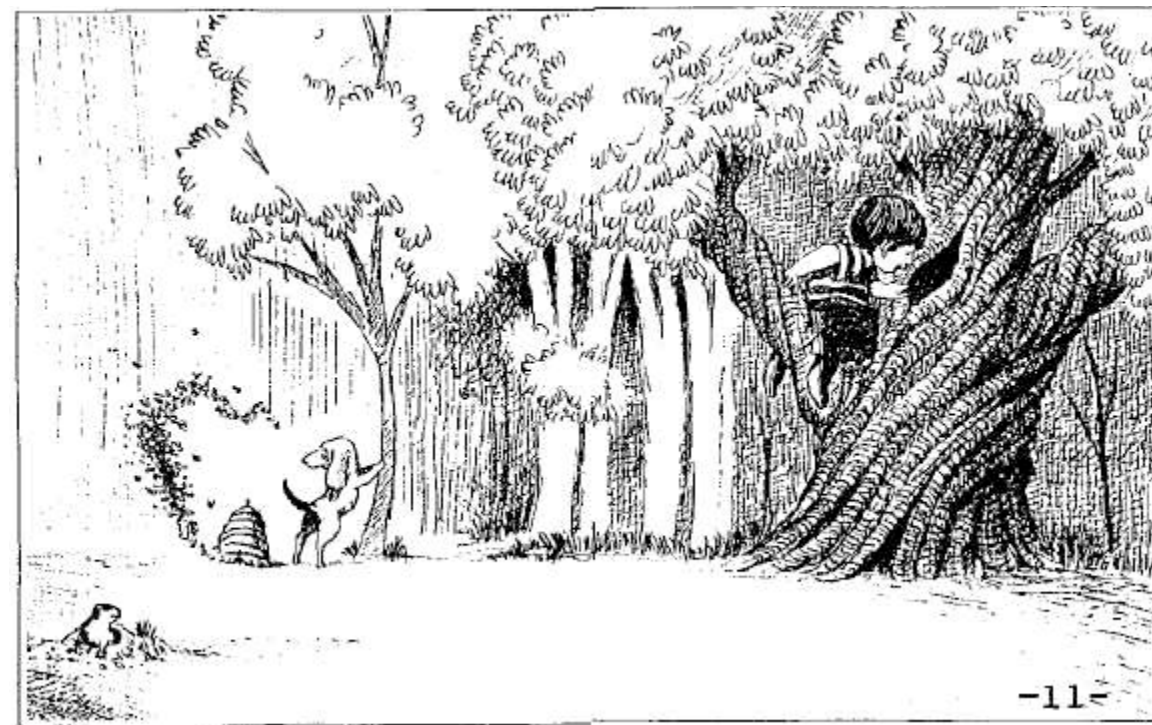
Pallotti, Gabriele. 2009. CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics* 30/4. pp. 490-601

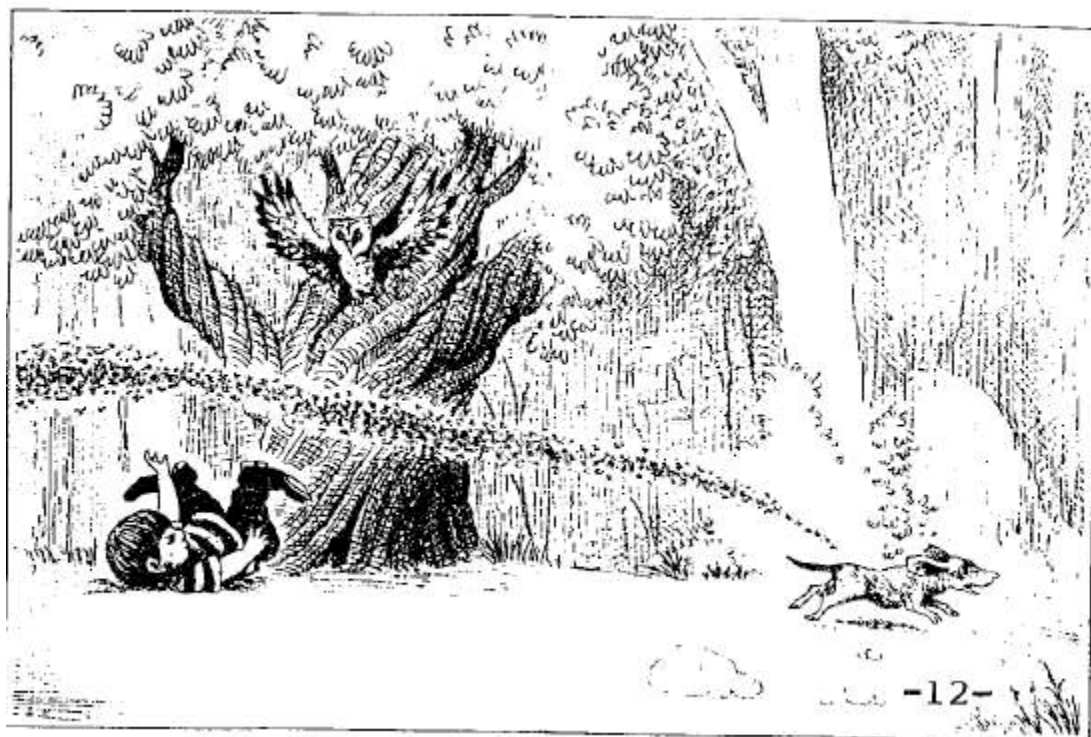
Reilly, J., Losh, M., Bellugi, U., Wulfeck, B. 2004. «Frog, where are you?» Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language* 88 pp. 229-247

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. *Relating events in narrative: Vol. 2. Typological and contextual perspectives* pp. 219-257. Mahwah, NJ: Lawrence Erlbaum Associates.

Toratani, Kiyoko. 2012. L2 Japanese Learners' Response to a Frog Story Elicitation Task. *Journal CAJLE*, Vol. 13. pp. 115-134.









Tillegg 2:

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina Karoliine Piksarv
(sünnikuupäev: 21.04.1984)

annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
**„Frosken, hvor er du?“ – syntaktisk kompleksitet i skriftlige fortellinger av estiske
innlærere av norsk som fremmedspråk.,**
(lõputöö pealkiri)

mille juhendaja on Anu Laanemets,
(juhendaja nimi)

reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas
digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja
lõppemiseni;

üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas
digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega
isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, _____ 15.08.2013

ERRATUM

Det mangler et avsnitt i kapittel 3, side 22 mellom de to siste avsnittene på siden
(*"Generell oversikt..."* og *"Bildeserien har blitt omfattende brukt..."*):

Tekstene baserer seg på en bildeserie fra en barnebok uten tekst - «Frog, where are you?» (Mayer 1969) - som består av 24 bilder (se tillegg 1) som sammen danner en historie av en liten gutt som leter etter en frosk som han har som kjæledyr. En dag finner gutten at frosken har forsvunnet og går sammen med sin hund ut for å lete etter den. Handlingsgangen finner sted i en rekke forskjellige steder og er ganske dynamisk, det opptrer også andre skogsdyr i «sideroller». Historien slutter med en bilde av gutten med en frosk i handen, men det er litt usikker hvilken frosk den er.